

【特許請求の範囲】

【請求項 1】

下記を具備する、画像を表す信号を損失無く圧縮し、符号化する方法：

損失のある圧縮されたデータファイルを発生する；

残差の圧縮されたデータファイルを発生する；および

損失のあるデータファイルを前記残差データファイルと結合し、損失の無いデータファイルを作成する、前記損失の無いデータファイルは、実質的にオリジナルデータファイルと同一である。

【請求項 2】

前記損失のある圧縮されたデータファイルおよび前記残差の圧縮されたデータファイルはフレーム内に基づいてまたはフレーム間に基づいて発生される、請求項 1 に記載の方法

10

【請求項 3】

発生は、ディスクリットコサイン変換（DCT）技術およびディスクリット四部木変換（DQT）技術を利用する、請求項 1 に記載の方法。

【請求項 4】

発生は、ゴロンブライズ符号化技術を利用する、請求項 1 に記載の方法。

【請求項 5】

下記を具備する、画像を損失無く圧縮し符号化する装置：

損失のある圧縮されたデータを発生する手段；

残差の圧縮されたデータファイルを発生する手段；および

前記損失のあるデータファイルと、前記残差データファイルを結合し、損失の無いデータファイルを作成する手段、前記損失の無いデータファイルは実質的にオリジナルデータファイルと同一である。

20

【請求項 6】

前記損失のある圧縮されたデータファイルを発生する手段および前記残差の圧縮されたデータファイルを発生する手段は、フレーム内に基づいて、またはフレーム間に基づいて発生される、請求項 5 に記載の装置。

【請求項 7】

前記発生手段は、ディスクリットコサイン変換（DCT）技術およびディスクリット四部木変換（DQT）技術を利用する、請求項 5 に記載の装置。

30

【請求項 8】

前記発生手段は、ゴロンブライズ技術を利用する、請求項 5 に記載の方法。

【請求項 9】

下記を具備する、画像を表す信号を損失無く、圧縮し、符号化する方法：

画像を表す信号を圧縮し、それにより、画像の圧縮されたバージョンを作成する；

前記画像の量子化された圧縮されたバージョンを順番に並べ、それにより、画像の順番に並べられた量子化された圧縮されたバージョンを作成する；

前記画像の圧縮されたバージョンを解凍する；

前記画像と前記画像の解凍されたバージョンとの間の差分を決定し、それにより、前記画像の残差バージョンを作成し；および

40

前記画像の損失のあるバージョンと前記画像の残差バージョンを出力し、前記画像の損失のあるバージョンと前記画像の残差バージョンとの組合せが実質的にオリジナル画像と同じである。

【請求項 10】

前記損失の無い圧縮はフレーム間に基づく、請求項 9 に記載の方法。

【請求項 11】

圧縮は、ディスクリットコサイン変換（DCT）技術およびディスクリット四部木（DQT）変換技術を利用する、請求項 9 に記載の方法。

【請求項 12】

50

順番に並べることはゴロンブーライズ符号化技術を利用する、請求項 9 記載の方法。

【請求項 13】

下記を具備する、複数のフレームから構成される画像情報を表す信号を損失無く圧縮し、符号化する方法：

第 1 フレームを圧縮し、それにより、前記画像の圧縮されたバージョンを作成する；

前記画像の圧縮されたバージョンを量子化し、それにより、前記画像の損失のあるバージョンを作成する；

前記画像の量子化された圧縮されたバージョンを順番に並べ、それにより、前記画像の順番に並べられた量子化された圧縮されたバージョンを作成する；

前記画像を表す信号の第 2 フレームを圧縮する；

前記画像の前記第 1 フレームと前記第 2 フレームとの間の差分を決定し、それにより、前記画像の残差バージョンを作成する；および

前記画像の残差バージョンを有した前記画像の損失のあるバージョンを出力する、前記画像の損失のあるバージョンと前記画像の残差バージョンとの組合せは実質的にオリジナル画像と同じである。

【請求項 14】

前記損失の無い圧縮はフレーム間に基づく、請求項 13 に記載の方法。

【請求項 15】

圧縮はディスクリットコサイン変換 (DCT) 技術およびディスクリット四部木変換 (DQT) 技術を利用する、請求項 13 に記載の方法。

【請求項 16】

順番に並べることはゴロンブーライズ符号化技術を利用する、請求項 13 に記載の方法。

【請求項 17】

画像を表す信号を損失無く圧縮し、符号化する方法：

画像を表す信号を圧縮し、それにより、画像の圧縮されたバージョンを作成する手段；

前記画像の圧縮されたバージョンを量子化し、それにより前記画像の損失のあるバージョンを作成する手段；

前記画像の前記量子化された圧縮されたバージョンを順番に並べ、それにより前記画像の順番に並べられた量子化された圧縮されたバージョンを作成する手段；

前記画像の圧縮されたバージョンを解凍する手段；

前記画像と前記画像の解凍されたバージョンとの間の差分を決定し、それにより、前記画像の残差バージョンを作成する手段；および

前記画像の損失のあるバージョンと前記画像の残差バージョンとを出力する手段、前記画像の損失のあるバージョンと前記画像の残差バージョンの組合せは実質的にオリジナル画像と同じである。

【請求項 18】

前記損失の無い圧縮はフレーム間に基づく、請求項 17 記載の装置。

【請求項 19】

前記圧縮手段はディスクリットコサイン変換 (DCT) 技術およびディスクリット四部木変換 (DQT) 技術を利用する、請求項 17 記載の装置。

【請求項 20】

20. 下記を具備する、画像情報を表す信号を損失無く圧縮し、符号化する装置：

第 1 フレームを圧縮し、それにより前記画像の圧縮バージョンを作成する手段；

前記画像の前記圧縮されたバージョンを量子化し、それにより前記画像の損失のあるバージョンを作成する手段；

前記画像の前記量子化された圧縮されたバージョンを順番に並べ、それにより、前記画像の順番に並べられた量子化された圧縮されたバージョンを作成する手段；

前記画像を表す信号の第 2 フレームを圧縮する手段；

前記画像の前記第 1 フレームと前記第 2 フレームとの間の差分を決定し、それにより、

10

20

30

40

50

前記画像の残差バージョンを作成する手段；および

前記画像の残差バージョンを有した前記画像の損失のあるバージョンを出力する手段、
前記画像の損失のあるバージョンと前記画像の残差バージョンとの組合せは実質的にオリジナル画像と同じである。

【請求項 21】

前記損失の無い圧縮は、フレーム間に基づく、請求項 20 に記載の装置。

【請求項 22】

前記圧縮する手段はディスクリットコサイン変換(DCT)技術およびディスクリット四部木変換(DQT)技術の組合せを利用する、請求項 20 に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像処理および圧縮に関する。特に、本発明は周波数領域のビデオ画像および音声情報の損失の無い符号化に関する。

【背景技術】

【0002】

デジタル画像処理は、デジタル信号処理の一般的な専門分野において、重要な位置を有する。人間の視覚認知の重要性は、デジタル画像処理の技術および科学において、著しい関心および進展を促進させた。フィルムまたは映画を投影するために使用されるビデオ信号のようなビデオ信号の送受信の分野において、種々の改良が画像圧縮技術になされつつある。現在および提案されたビデオシステムの多くは、デジタル符号化技術を利用する。この分野の観点は、画像符号化、画像復元および画像特徴選択を含む。画像符号化は、効率的な方法でデジタル通信の画像を送信するための試みを表し、同時に、必要とされる帯域幅をできるだけ最小にするように数ビットを使用し、ある制限内で歪みを維持する。画像復元は物体の真の画像を再生するための努力をあらわす。通信チャネル上に送信される符号化された画像は種々の要因により歪みを生じるかもしれない。質が低下する原因は、もともと物体から画像を創出する際に生じていたかもしれない。特徴選択は、画像のある属性の選択に言及する。そのような属性は、より広い文脈において、認識、分類および決定において必要かもしれない。

【0003】

デジタル映画のようなビデオのデジタル符号化は、改良された画像圧縮技術から利益を得る領域である。デジタル画像圧縮は、一般的に、2つのカテゴリに分類されるかもしれない。すなわち、損失の無い方法と損失のある方法である。損失の無い画像は、いかなる情報の損失も無しに再生される。損失のある方法は、以下を含む：圧縮比、圧縮アルゴリズムの質、およびアルゴリズムの実施に応じて多少の情報の回復できない損失。一般に、費用効率が低いデジタル映画手法のために要望される圧縮比を得るために損失のある圧縮手法が考慮される。デジタル映画の品質レベルを得るために、圧縮手法は、視覚的に損失の無いレベルの性能を提供しなければならない。従って、圧縮処理の結果として情報の数学的な損失があるけれども、この損失によって生じた画像歪みは通常の画像観察下で視聴者に感知できないようであればならない。

【0004】

既存のデジタル画像圧縮技術は、ほかの応用のために開発されてきた、たすなわち、テレビシステムのために開発されてきた。そのような技術は、意図したアプリケーションに対して設計上の妥協を適当なものとしたが映画プレゼンテーションに必要な品質要件に適合しない。

【0005】

デジタル映画圧縮技術は、映画ファンが以前経験した視覚の質を提供しなければならない。理想的には、デジタル映画の視覚品質は、高品質の公開版フィルムの品質を超えようと試みなければならない。同時に、この圧縮技術は高符号化効率を実用的にさせなければならない。本願明細書において定義されるように、コーディング効率はある質のレベルを

満たすために圧縮された画像品質に必要なビットレートに言及する。さらに、システムおよび符号化技術は、異なるフォーマットに対応するように内蔵の柔軟性を有さなければならない、費用効率が高くなければならない。すなわち、小さなサイズで効率的なデコーダまたはエンコーダ処理でなければならない。

【0006】

入手可能な多くの圧縮技術は、有意水準の圧縮を提供するが、結果として、ビデオ信号の品質の劣化を生じる。典型的に、圧縮された情報を転送するための技術は、圧縮された情報が一定のビットレートで転送されることを必要とする。

【0007】

ビデオ信号に対して所望のレベルの品質を維持しながら、有意水準の圧縮を提供することができる1つの圧縮技術は、符号化された離散コサイン変換(DCT)係数データの順応して大きさに合わせて作られたブロックおよびサブブロックを利用する。この技術は、以下において、順応ブロックサイズ離散コサイン変換(ABSDCT)方法と呼ばれるであろう。この技術は、この発明の譲受人に譲渡され、参照することにより、本願明細書に組み込まれる米国特許番号第5,021,891号、(発明の名称:「順応ブロックサイズ画像圧縮方法およびシステム」)に開示される。DCT技術は、また、この発明の譲受人に譲渡され参照することにより本願明細書に組み込まれる米国特許番号第5,107,345号(発明の名称:「順応ブロックサイズ圧縮方法およびシステム」)にも開示される。さらに、微分四部木と組み合わせたABSDCT技術も、この発明の譲受人に譲渡され、参照することにより、本願明細書に組み込まれる、米国特許番号第5,452,104号(発明の名称:「順応ブロックサイズ画像圧縮方法およびシステム」)に開示されている。これらの特許において開示されるシステムは、「フレーム内」符号化と呼ばれるものを利用し、画像データの各フレームはその他のフレームの内容に関係なく符号化される。ABSDCT技術を用いて、なし逃げられるデータレートは、画質の識別可能な劣化のない約15億ビット/秒からおよそ5000万ビット/秒まで、低減してもよい。

【0008】

ABSDCT技術は、白黒カラー画像を圧縮するかまたは画像を表す信号を圧縮するために使用してもよい。カラー入力信号は、YIQフォーマットであってもよく、Yは輝度または光度サンプルであり、IおよびQは、各4:4:4または代替フォーマットに対して、クロミナンスまたはカラーサンプルである。YUV、YCbCrまたはRGBフォーマットのような他の周知のフォーマットも、使用してもよい。色に対する目の低空間の感度のため、大部分の調査は、色成分の副サンプルが水平および垂直方向に4だけ合理的であることを示した。したがって、ビデオ信号は4つの輝度サンプルおよび2つのクロミナンス・サンプルによって表されることができる。

【0009】

ABSDCTを用いて、ビデオ信号は一般に処理のため、ピクセルのブロックに分割されるであろう。各々のブロックに対して、輝度およびクロミナンス成分は、ブロックサイズ割当てエレメントまたはブロックインターリーバーに渡される。例えば、 16×16 (ピクセル)ブロックは、ブロックインターリーバーに渡してもよい。ブロックインターリーバーは各 16×16 ブロック内で画像サンプルを順序付けるかまたは組織化して離散コサイン変換(DCT)解析のためにデータのブロックおよび合成副ブロックを生成する。DCTオペレータは、時間および空間のサンプル信号を同じ信号の周波数表示に変換する1つの方法である。周波数表示に変換することによって、DCT技術は、量子化器が画像の周波数分布特性を利用するように設計することができるように、非常の高レベルの圧縮を可能にすることを示してきた。好適実施形態において、1つの 16×16 DCTは第1の順序付けに適用され、4つの 8×8 DCTは第2の順序付けに適用され、16の 4×4 DCTは、第3の順序付けに適用され、および64の 2×2 DCTは第4の順序付けに適用される。

【0010】

DCT操作は、ビデオソースに固有の空間冗長度を低減する。DCTが実行された後、殆どのビデオ信号エネルギーは、2乃至3のDCT係数に集中する傾向がある。さらなる

10

20

30

40

50

変換、微分四部木変換(DQT)は、DCT係数の中で冗長度を低減するために使用してもよい。

【0011】

16×16のブロックおよび各副ブロックに対して、DCT係数値および(DQTが使用されるなら)DQT値が解析され、ブロックまたは副ブロックを符号化するために必要なビット数を決定する。次に、符号化するために最少数のビットを必要とする副ブロックの組合せが選択され、画像セグメントを表す。例えば、2つの8×8副ブロック、6つの4×4副ブロックおよび8つの2×2副ブロックは、画像セグメントを表すために選択してもよい。

【0012】

次に、選択されたブロックまたは副ブロックの組合せは適切に順番に16×16のブロックに配列される。次に、DCT/DQT係数値は、送信の準備のために、周波数重みづけられ、量子化され、(可変長符号化)のような符号化をしてもよい。上記したABSDCT技術は、著しくよく実行するけれども、それはコンピュータを必要とするほど強力である。

【0013】

さらに、上述したABSDCTの使用が視覚的に損失が無いけれども、時には、データが符号化される正確な方法でデータを再生することが望ましい。例えば、マスタリングおよび永久保存の目的は、正確にものとの状態に再生することができるようにデータを圧縮する必要がある。

【0014】

典型的に損失の無い画像のための圧縮システムは、予測子から構成され、予測子は符号化される現在のピクセルの値を推定する。残りのピクセルは実際のピクセルと予測されるピクセルとの間の差分として得られる。次に、残りのピクセルはエントロピー符号化され記憶されたものは送信される。予測はピクセルの相関を除去するので、残りのピクセルは特徴のある両面のある指数(ラブラシアン)分布を有した低減されたダイナミックレンジを有する。それゆえに、圧縮を有する。残りの圧縮量は、予測および次のエントロピー符号化方法に依存する。大部分の共通して使う予測方法は、差分PCM(DPCM)および適応DPCM(ADPCM)のようなその変形である。

【発明の開示】

【発明が解決しようとする課題】

【0015】

ベルに基づく予測を伴う問題は、残差がまだ高エネルギーを有するという点である。それは、少数の隣接したピクセルだけが予測方法において使われるという事実による。したがって、ベルに基づく予測スキームの符号化効率を改良する余地がある。

【課題を解決するための手段】

【0016】

本発明の実施例は、圧縮を達成するために、損失の無い方法でデジタル画像およびビデオデータを圧縮するためのシステムを記載する。システムは、ハイブリッドである。このことは、損失のある方法で前記データを圧縮する一部と損失の無い態様で残差データを圧縮する一部とを有することを意味する。損失のある部分に対して、システムは適応ブロックサイズ離散コサイン変換(ABSDCT)アルゴリズムを使用する。ABSDCTシステムは、高い表示品位および圧縮比を生じる前記データを圧縮する。残差は、オリジナルと、ABSDCTシステムから解凍されたものとの間の差分として得られる。この残差は、ゴロンブ・ライス(Golomb-Rice)符号化アルゴリズムを用いて損失無く符号化される。視覚的に基づいた適応ブロックサイズおよびDCT係数の量子化により、残差は非常に低いエネルギーを有し、従って、全体に損失の無い圧縮比を生じる。

【0017】

ABSDCTシステムは、映画の品質において、高い圧縮比を得る。ABSDCTシステムはブロックに基づくので、ABSDCTシステムはいかなるベルに基づくスキームよ

10

20

30

40

50

りも非常によいピクセル相関を除去する。したがって、ABSDCTシステムは本願明細書において記載される損失の無いシステムにおける予測として使用される。この予測子と連動して、損失の無い符号化システムが加えられ、ハイブリッドの損失の無い圧縮システムを形成する。留意すべきは、システムは、動画像同様に静止画像を圧縮することができるという点である。静止画像なら、ABSDCT圧縮データおよびエントロピ符号化された残差データのみが圧縮された出力として使用される。運動順序のために、フレーム内圧縮を使用するかまたはフレーム間圧縮を使用するかの決定がなされる。例えば、 $f(t)$ が時刻 t における画像フレームを表すなら、 $F(t)$ および $F(t + \Delta t)$ は、それぞれ時刻 t および $t + \Delta t$ における画像フレームのDCTを示す。 Δt が2つの連続的なフレーム間の時間間隔と対応する点に注意する必要がある。

【0018】

この発明は、データが符号化された正確な方法でデータを再生することを可能にするデータを圧縮するための装置および方法に具現化される。実施形態は、フレーム内符号化、フレーム間符号化または、それら2つのハイブリッドを実行するシステムを含む。このシステムは、離散コサイン変換係数データの順応した大きさに合わせて作られたブロックおよび副ブロックを利用する品質に基づいたシステムである。1つのブロックのピクセルデータは、エンコーダに対する入力である。エンコーダは、ブロックサイズ割当て(BSA)構成要素を含み、ブロック割当て構成要素は処理のためにピクセルの入力ブロックを分割する。ブロックサイズ割当ては、入力ブロックおよびさらに再分割されたブロックの分散に基づく。一般に、ブロックおよび副ブロックの平均値が異なる所定の範囲に入るなら、より大きい分散を伴うエリアは、より小さなブロックに再分割され、より小さな分散を伴うエリアは再分割されない。従って、最初に、ブロックの分散閾値は、その平均値に応じて公称値から変更され、次に、ブロックの分散はこの閾値と比較され、そして、分散が閾値より大きい場合、そのブロックは再分割される。

【0019】

ブロックサイズ割当ては変換構成要素に供給され、変換構成要素はピクセルデータを周波数領域データに変換する。この変換は、ブロックサイズ割当てを介して選択されたブロックおよび副ブロックのみに実行される。AC構成要素の場合、量子化および連続化を介してその変換データにスケールリング処理をする。変換データの量子化は、コントラスト、係数カウント、レート歪み、ブロックサイズ割当ての密度及び/または過去スケールファクタに対して調節するスケールファクタのような、画像品質測定基準に基づいて量子化される。ジグザグスキャンのような連続化は同じ値の最も長い可能なランレングスを作ることに基づく。次に、データのストリームは送信に備えて、可変長符号化器により符号化される。コーディングはハフマン・コーディングであってもよい、または、ゴロンブ・ライス符号化のような指数分布に基づいてもよい。

【0020】

ABSDCTのようなハイブリッド圧縮システムの使用は、ピクセルまたはDCT値の良好な予測子のように動作する。それゆえ、ABSDCTは、ペルに基づく予測を使用しているシステムより高い損失の無い圧縮比を生じる。損失のある部分は、デジタル映画品質結果を供給する。すなわち、視覚的に損失の無いファイルに圧縮結果を供給する。損失の無い部分の場合、ハフマンコードと異なり、ゴロンブ・ライス符号化はいかなる先験的符号発生を必要としない。それゆえ、ゴロンブ・ライス符号化は、ハフマン符号化の場合に記憶される拡張コードブックを必要としない。これは、チップの有効スペースの効率的な使用を生じる。それゆえに、チップ・サイズは、ハードウェア実施において低減される。さらに、ゴロンブ・ライス符号化は、ハフマン符号化より実施するのにより簡単である。また、DCT係数または残りのピクセルは必然的に指数分布を有するので、ゴロンブ・ライス符号化は、ハフマン符号化よりも高い符号化効率を得る。さらに、圧縮システムの損失のある部分がブロック再分割において視覚的に重要な情報を使用するので、コンテキストモデリングは残差符号化において固有である。これは、残差符号化のためにコンテキストデータを集めるのに、余分な記憶レジスタが必要ないという点で、重要である。ま

た、運動推定が使われないので、システムは実施するのに非常に簡単である。

【0021】

画像情報を表す信号を損失が無いように圧縮し符号化するための装置および方法が請求される。画像情報を表す信号は、圧縮され、画像の圧縮されたバージョンを作る。画像の圧縮されたバージョンは量子化され、それにより、画像の損失のあるバージョンを作る。画像の圧縮されたバージョンは、また、順番に並べられ、画像の順番に並べられた量子化された圧縮されたバージョンを作る。このバージョンは、次に、解凍され、オリジナル画像と解凍されたバージョンとの間の差分が決定され、それにより画像の残差バージョンを作る。画像の損失のあるバージョンおよび画像の残差バージョンは、別個にまたは結合されて別に出力してもよく、その場合、画像の解凍された損失のあるバージョンと、画像の残差バージョンとの組合せは実質的にオリジナル画像と同じである。

10

【0022】

画像情報を表す信号を損失無く圧縮し、符号化する方法が、請求される。損失のある圧縮されたデータファイルおよび残りの圧縮されたデータファイルが発生される。損失のある圧縮されたデータファイルおよび残りの圧縮されたデータファイルが結合されると、実質的にオリジナルデータファイルと同一であり損失の無いデータファイルが作られる。

【0023】

したがって、効率的に損失の無い圧縮を提供するための装置および方法を提供することが、実施形態の1つの観点である。

【0024】

マスタリングおよびアーカイブの目的にの助けとなる方法で、損失無く、デジタル画像情報およびオーディオ情報を圧縮することは実施形態の他の観点である。

20

【0025】

フレーム間に基づいて、損失の無い圧縮システムを提供することは実施形態の他の観点である。

【0026】

フレーム内に基づいて損失の無い圧縮システムを提供することは実施形態の他の観点である。

【発明を実施するための最良の形態】

【0027】

デジタル信号のデジタル伝送を容易にして、対応する利益を享受するために、一般的にある形態の信号圧縮を採用することが必要である。結果として生じる画像の高い圧縮を実現する一方、画像の高品質が維持されることもまた、重要である。さらに、コンパクトなハードウェア実施のための計算効率が望まれる。この計算効率は多くのアプリケーションにおいて重要である。

30

【0028】

本発明の一実施形態を詳細に説明する前に、本発明は、そのアプリケーションにおいて、構成の詳細および以下の記載で述べるまたは、図で図解される部品の配列に限定されないことが理解されるべきである。本発明は他の実施形態を実施することができ、種々の方法で実行される。また、ここで使用される言葉遣いおよび専門用語は、記載の目的のためであり、制限と見なされるべきではない。

40

【0029】

一実施形態の一観点において採用される画像圧縮は、この発明の譲受人に譲渡され、参照することによりここに組み込まれる、1999年11月8日に出願されたシリアル番号第09/436,085の同時係属米国特許出願（発明の名称：「コントラスト感応分散に基づく適応ブロックサイズ画像圧縮」）に開示されるような離散コサイン変換（DCT）に基づく。DCTを利用した画像圧縮および解凍システムは、この発明の譲受人に譲渡され、参照することによりここに組み込まれる、2000年、1月28日に開示されたシリアル番号第09/494,192号の同時係属米国特許出願（発明の名称：「品質に基づく画像圧縮」）に記載されている。一般に、デジタル領域において処理される画像は、

50

$N \times N$ のサイズの重畳しないブロックの配列に分割されたピクセルデータから成る。二次元のDCTを、各ブロック上で実行してもよい。二次元のDCTは、以下の関係によって定義される：

【数1】

$$X(k, l) = \frac{\alpha(k)\beta(l)}{\sqrt{N \cdot M}} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(m, n) \cos\left[\frac{(2m+1)\pi k}{2N}\right] \cos\left[\frac{(2n+1)\pi l}{2N}\right]$$

$$, 0 \leq k, l \leq N-1$$

10

【0030】

但し、

【数2】

$$\alpha(k), \beta(k) = \begin{cases} 1, & \text{if } k = 0 \\ \sqrt{2}, & \text{if } k \neq 0 \end{cases}$$

【0031】

および $x(m, n)$ は $N \times M$ ブロック内のロケーション (m, n) におけるピクセルであり $X(k, l)$ は対応するDCT係数である。

【0032】

ピクセル値が0以上であるので、DCT成分 $X(0, 0)$ は常に正で、通常最も多くのエネルギーを有する。実際、代表的な画像の場合、大部分の変換エネルギーは、成分 $(0, 0)$ 周辺に集中する。このエネルギー圧縮特性は、DCT特性をかくも魅力的な圧縮方法にしているところのものである。

【0033】

画像圧縮技術は、コントラスト適応符号化を利用して、さらなるビットレートの低減を得る。大部分の自然の画像は、相対的にゆっくりと変化する平らな領域、および物体境界および高コントラストテクスチャのような人通りの多い場所から構成される。コントラスト適応符号化スキームは、より多くのビットを人通りの多い場所に割当てより少ないビットをより少ない人通りの多い場所に割当てることにより、この要因を利用する。

【0034】

コントラスト適応方法は、フレーム間符号化（時空間的処理）の代わりに、フレーム内符号化（空間処理）を利用する。フレーム間符号化は、本質的に、より多くの複雑の処理回路に加えて複数のフレームバッファを必要とする。多くのアプリケーションにおいて、複雑さを低減することは、実際の実施のために必要である。フレーム間符号化は、また時空間的符号化スキームの機能が衰えて、その動作が悪化した状況で有効である。例えば、毎秒24コマの映画は、機械的シャッターによる積分時間が相対的に短いので、このカテゴリに入れることができる。短い積分時間は、より高い度合いの時間的エリアシングを可能にする。高速に移動すると完結的になるので、フレームとフレームの相関関係の仮定がこわれる。フレーム内符号化はまた、50Hzおよび60Hzの電力線周波数が含まれる時に、標準化するのがより容易である。テレビ放送は、現在50Hzか60Hzで信号を送る。デジタル手法である、フレーム内スキームの使用は50Hzおよび60Hzの両方に適合する。あるいは、空間分解能に対してフレームレートをトレードオフすることにより毎秒24コマの映画にさえも適用できる。

【0035】

画像処理の目的のために、重畳しないブロックの配列に分割されるピクセルデータにDCT動作が実行される。ブロックサイズは $N \times N$ のサイズであるとして本願明細書におい

20

30

40

50

て論じられるけれども、種々のブロックサイズを用いても良いことが想像される点に留意すべきである。例えば、 $N \times M$ のブロックサイズは、 M が N より大きいまたは未満であって、 N および M の両方が整数である場合に、利用してもよい。他の重要な観点は、ブロックが、 $N/i \times N/j$ 、 $N/i \times M/j$ 、 $N/i \times M/j$ (但し、 i および j は整数)等のような少なくとも1つのレベルの副ブロックに分割可能である。さらに、本願明細書において議論される例示ブロックサイズは、DCT係数の対応するブロックおよび副ブロックを有する 16×16 のピクセルブロックである。偶数または奇数の両方の整数値、例えば 9×9 のような種々他の整数値を使用しても良いことがさらに想像される。

【0036】

図1および図2は構成可能なシリアルライザの概念を組み込んだ画像処理システム100を図解する。画像処理システム100は受信したビデオ信号を圧縮するエンコーダ104を含む。圧縮信号は、伝送路または物理的な媒体108を使用して送られ、デコーダ112によって受信される。デコーダ112は、受信した符号化されたデータを画像サンプルに復号し、次に、それを表示してもよい。

【0037】

一般に、画像は処理のためにピクセルのブロックに分割される。イメージは、加工度のための面素のブロックに分割される。カラー信号は、RGB-YC1C2変換器116を使用して、RGB空間からYC1C2空間に変換してもよい。この場合、Yは輝度成分、または光度成分であり、C1およびC2はクロミナンス成分または色成分である。色に対する眼の空間感度が低いので、多くのシステムは水平方向および垂直方向にC1成分およびC2成分を4だけサブサンプリングする。しかしながら、サブサンプリングは必要でない。4:4:4フォーマットとして知られる、完全分解能はデジタル映画をカバーするものと言われているようないくつかのアプリケーションにおいて非常に有効または必要かもしれない。2つの可能なYC1C2表示はYIQ表示およびYUV表示であり、両方は技術的によく知られている。また、YCbCrとして知られているYUV表示の変化を採用することも可能である。これを、さらに、奇数成分と偶数成分に分解してもよい。したがって、実施の形態において、表示Y-偶数、Y-奇数、Cb-偶数、Cb-奇数、Cr-偶数、Cr-奇数が使用される。

【0038】

好適実施形態において、偶数および奇数のY、Cb、およびCr成分の各々はサブサンプリング無しに処理される。このように、 16×16 ブロックのピクセルの6つの成分の各々の入力はエンコーダ104に供給される。図示するために、Y偶数成分のためのエンコーダ104が図解されている。同様のエンコーダが、Y-奇数成分および偶数および奇数のCbおよびCr成分のために使用される。エンコーダ104は、ブロックサイズ割当て構成要素120を含み、ブロックサイズ割当て構成要素120はビデオ圧縮の準備のために、ブロックサイズ割当てを実行する。ブロックサイズ割当て構成要素120は、ブロック内の画像の知覚特性に基づいて 16×16 のブロックのブロック分解を決定する。ブロックサイズ割当ては、各 16×16 のブロックを、 16×16 ブロック内の行動に応じて、四部木態様で 8×8 、 4×4 、および 2×2 のようなより小さなブロックにさらに分割する。ブロックサイズ割当て構成要素120は、長さが1ビットと2ビットとの間にあるPQRデータと呼ばれる、四部木データを発生する。従って、ブロックサイズ割当てが、 16×16 ブロックサイズを分割すると決定するなら、PQRデータのRビットが設定され、4つの分割された 8×8 ブロックに対応するQデータの4つのさらなるビットが続く。ブロックサイズ割当てが、 8×8 ブロックのいずれかを分割することを決定するなら、さらに分割された 8×8 ブロックの各々に対してPデータのさらなるビットが追加される。

【0039】

図3を参照すると、ブロックサイズ割当て構成要素120の動作の詳細を示すフロー図が提供される。ブロックの分散が、ブロックをさらに分割するという決定の測定基準として使われる。ステップ202から始まって、 16×16 ブロックのピクセルが読まれる。

ステップ204において、 16×16 のブロックの分散が計算される。

【0040】

分散は、次のように計算される：

【数3】

$$\text{var} = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_{i,j}^2 - \left(\frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_{i,j} \right)^2$$

10

【0041】

但し、 $N=16$ であり、 $x_{i,j}$ は $N \times N$ ブロック内の i 行 j 列のピクセルである。ステップ206において、閾値 T_{16} の分散が変更され新しい閾値 T'_{16} が与えられ、ブロックの平均値が2つの所定値間にあるなら、そのブロック分散は新しい閾値 T'_{16} と比較される。

【0042】

分散 v_{16} が閾値 T_{16} より大きくなければ、ステップ208において、 16×16 ブロックの開始アドレスが一時記憶装置に書き込まれ、PQRデータのRビットは0に設定され、 16×16 ブロックの再分割されないことを示す。次に、アルゴリズムは次の 16×16 ブロックのピクセルを読む。分散 v_{16} が閾値 T_{16} より大きければ、ステップ210において、PQRデータのRビットが1に設定され、 16×16 ブロックが4つの 8×8 ブロックに再分割されることを示す。

20

【0043】

4つの 8×8 ブロック、すなわち $i=1:4$ は、ステップ212に示すように、さらなる再分割のために順次考慮される。各 8×8 ブロックに対して、分散、 v_{8i} がステップ214において、計算される。ステップ216において、最初に、分散閾値 T_8 が変更され、新しい閾値 T'_8 が与えられ、ブロックの値が2つの所定値間にあれば、ブロック分散はこの新しい閾値と比較される。

【0044】

30

ステップ218において、分散 v_{8i} が閾値 T_8 より大きくなければ、 8×8 ブロックの開始アドレスは一時記憶装置に書き込まれ、対応するQビット、 Q_i は0に設定される。次に、次の 8×8 ブロックが処理される。ステップ220において、分散閾値 T_8 より大きければ、対応するQビット、 Q_i は1に設定され、 8×8 ブロックが4つの 4×4 ブロックに再分割されることを示す。

【0045】

4つの 4×4 ブロック $j_i=1:4$ は、ステップ222に示すように、さらなる再分割のために、順次考慮される。各 4×4 ブロックに対して、分散、 v_{4ij} はステップ224において、計算される。ステップ226において、最初に、分散閾値 T_4 が変更され、新しい閾値 T'_4 が与えられ、ブロックの平均値が2つの所定間であれば、ブロック分散はこの新しい閾値と比較される。

40

【0046】

分散 v_{4ij} が閾値 T_4 より大きくなければ、ステップ228において、 4×4 ブロックのアドレスが書き込まれ、対応するPビット、 P_i は0に設定される。次に、次の 4×4 ブロックが処理される。分散 v_{4ij} が閾値 T_4 より大きければ、ステップ230において、対応するPビット、 P_{ij} は1に設定され、 4×4 ブロックは4つの 2×2 ブロックに再分割されることを示す。さらに、4つの 2×2 ブロックのアドレスは一時記憶装置に書き込まれる。

【0047】

閾値 T_{16} 、 T_8 および T_4 は所定の定数であってよい。これは、つらい決断として知

50

られる。あるいは、適応的意思決定または柔軟な意思決定を実施してもよい。例えば、柔軟な意思決定は 2×2 ブロックの平均ピクセル値に応じて分散のための閾値を変化させる（但し N は8、4、または2の値を取り得る）。従って、平均ピクセル値の閾値を閾値として用いてもよい。

【0048】

説明のために以下の例を考察する。Y成分のための所定の分散閾値を、それぞれ 16×16 ブロック、 8×8 ブロック、および 4×4 ブロックの場合に、50、1100、および880とする。言い換えれば、 $T16 = 50$ 、 $T8 = 1100$ 、および $T4 = 880$ とする。平均値の範囲を80および100とする。 16×16 ブロックの計算された分散を60と仮定する。60は $T16$ より大きいので、平均値90は80と100との間にあり、 16×16 ブロックは4つの 8×8 の副ブロックに再分割される。 8×8 ブロックの計算された分散が、1180、935、980、および1210であると仮定する。 8×8 ブロックのうちの2つは $T8$ を越える分散を有するので、これらの2つのブロックはさらに再分割され、合計8つの 4×4 の副ブロックを生成する。最後に、8つの 4×4 ブロックの分散が620、630、670、610、590、525、および690であり、対応する平均値が90、120、110、115であると仮定する。最初の 4×4 ブロックの平均値は範囲（80、100）に入るので、閾値は、880未満である $T'4 = 200$ に下がるであろう。従って、この 4×4 ブロックは7番めの 4×4 ブロックと同様に再分割される。

【0049】

同様の手続を用いて、輝度成分Y一奇数および色成分CbおよびCrに対してブロックサイズを割当てる点に留意する必要がある。色成分は、水平方向、垂直方向、またはその両方において、大幅に低減してもよい。

【0050】

さらに、ブロックサイズ割当てがトップダウン手法として記述されたけれども、最大のブロック（この例では 16×16 ）が最初に評価され、その代わりにボトムアップ手法を用いても良いことに留意する必要がある。ボトムアップ手法は、最初に最も小さいブロック（この例では、 2×2 ）を評価するであろう。

【0051】

図1に戻ると、PQRデータは選択されたブロックのアドレスとともに、DCT構成要素124に供給される。DCT構成要素124はPQRデータを用いて、選択されたブロックに関して適当なサイズの離散コサイン変換を実行する。選択されたブロックだけがDCT処理を行う必要がある。

【0052】

画像処理システム100はまた、DCTのDC係数の中で冗長度を低減するためにDQT構成要素128から成る。DC係数は、各DCTブロックの左上隅にある。DC係数は、一般に、AC係数に比べて大きい。サイズにおける不一致は、有効な可変長コーダを設計することを困難にさせる。従って、DC係数の中で冗長度を低減することは利点がある。

【0053】

DQT構成要素128はDC係数に2-D DCTsを実行し、一度に 2×2 の単位で行なう。 4×4 ブロック内で、 2×2 ブロックから始めて、4つのDC係数に関して2-D DCTが実行される。この 2×2 DCTは微分四部木変換、または4つのDC係数のDQTと呼ばれる。次に、 8×8 ブロック内の3つの隣接するDC係数とともにDQTのDC係数を用いて次のレベルDQTを計算する。最後に、 16×16 ブロック内の4つの 8×8 ブロックのDC係数を用いてDQTを計算する。従って、 16×16 ブロックにおいて、1つのDC係数があり、残りはDCT及びDQTに対応するAC係数である。

【0054】

変換係数（DCTおよびDQTの両方）は量子化のために量子化器に供給される。好適実施形態において、DCT係数は、周波数重み付けするマスク（FWMs）および量子化

10

20

30

40

50

スケール因子を用いて量子化される。FWMは、入力DCT係数のブロックと同じ次元の周波数重みのテーブルである。周波数重みは、異なるDCT係数に異なる重みを印加する。重みは、人間の視覚または光学システムがより敏感である周波数内容を有する入力サンプルを強調するように設計され、そして、視覚または光学システムがより敏感でない周波数内容をされている。重みは、また見える距離等のような因子に基づいて設計してもよい。

【0055】

重みは、経験によるデータに基づいて選択される。8×8DCT係数のための重み付けマスクを設計するための方法は、ISO/IEC JTC1 CD 10918「連続階調静止画像のデジタル圧縮および符号化—パート1：要件およびガイドライン」、国際標準化機構1994、これは参照することにより本願明細書に組み込まれる。一般に、2つのFWMは、1つが輝度成分のためのそして1つがクロミナンス成分のために設計される。ブロックサイズ2×2、4×4のためのFWMテーブルは大幅に減らすことにより得られ、16×16のためのFWMテーブルは、8×8ブロックのためのFWMテーブルの補間により得られる。スケールファクタは、量子化された係数の質およびビットレートを制御する。

【0056】

従って、各DCT係数は以下の関係式に従って量子化される。

【数4】

$$DCT_q(i, j) = \left\lfloor \frac{8 * DCT(i, j)}{fwm(i, j) * q} \pm \frac{1}{2} \right\rfloor$$

但し、DCT(i, j)は入力DCT係数であり、fwm(i, j)は周波数重み付けするマスクであり、qはスケールファクタであり、およびDCT_q(i, j)は量子化された係数である。DCT係数の符号に応じて、大括弧内の最初の項は端数のない形に切り上げるかまたは端数が切り捨てられることに留意する必要がある。DQT係数は、また、適切な重み付けマスクを使用して量子化される。しかしながら、複数のテーブルまたはマスクを使用することができ、Y成分、Cb成分、およびCr成分の各々に印加することができる。

【0057】

次に、AC値がDC値から分離され130、別個に処理される。DC構成要素に対して、各スライスの最初のDC成分値が符号化される。従って、各スライスの各次のDC成分値は、自分自身と自分より先行するDC成分値との差分として表され、符号化される134。損失の無い符号化の場合、各スライスの初期DC成分値と前記差分は図6および図8に対して記載したように、ゴロンブーライスをを用いて符号化される138。連続するDC成分値間の差分のためにゴロンブーライスを符号化を使用することは、DC成分値の差が2辺のある指数分布を有する傾向があるという点で利点がある。次にデータは、一時的にバッファ142を用いて記憶され、次に、送信チャネル108を介してデコーダ112に転送または送信される。

【0058】

図8はDC成分値を符号化するプロセスを図解する。このプロセスは等価的に静止画、（これらに限定されるわけではないが、映画または高品質テレビジョンのような）ビデオ画像、およびオーディオに適用できる。データ804の所定のスライスに対して、スライスの第1のDC成分値が検索される808。次に、第1のDC成分値が符号化される812。AC成分値と異なって、DC成分値は、量子化される必要はない。一実施形態において、16×16ブロックのための単一DC値は、ブロックサイズの割当てが停止しているにもかかわらず、使用される。8×8または4×4、またはブロックサイズ割当てにより定義されるいずれかの可変ブロックサイズのような異なる固定のサイズのブロックも使用できると予想される。次に、所定のスライスの第2または次の、DC成分値が検索され

る 816。次に第 2 の D C 成分値が第 1 の D C 成分値と比較され、その差分すなわち残差が符号化される 820。従って、第 2 の D C 成分値は、自分自身と第 1 の値との間の差分として表すしか必要としない。このプロセスは、スライスの各 D C 成分値に対して繰り返される。このように、問合せ 824 はスライス（最後のブロックそれゆえ、最後の D C 値）の終わりに到達したかどうかに関してなされる。そうでなければ 828、スライスの次の D C 値が検索され 816、プロセスは繰返される。そうであれば、832、次のスライスが検索され 804、このプロセスは、フレームの全てのスライス、およびファイルの全てのフレームが処理されるまで繰返される。

【0059】

D C 成分値の損失の無い符号化の目的は、低分散を有する傾向がある残差値を生成することである。DCT を使用する際に、D C 係数成分値は、最大の画素エネルギーを与える。それゆえ、D C 成分値を量子化しないことによって、残差の分散は、低減される。

【0060】

A C 要素の場合、データのブロックおよび周波数重み付けマスクは次に、量子化器 146 またはスケールファクターエレメントにより倍率がかけられる。D C T 係数の量子化は、大多数の D C T 係数をゼロまで低減し、その結果圧縮が生じる。好適実施形態において、平均ビットレートに対応する 32 のスケールファクタがある。MPEG2 のような他の圧縮方法と異なり、平均ビットレートは目標のビットレートおよびバッファステータスの代わりに処理される画像の品質に基づいて制御される。

【0061】

さらに圧縮を増大させるために、量子化された A C 係数はシリアライザ 150 に供給される。シリアライザ 150 は量子化された係数のブロックを走査し、順番に並べられたストリームの量子化された係数を生成する。ジグザグ走査、カラム走査または行走査を採用してもよい。ジグザグ以外のパターン同様、多数の異なるジグザグ走査パターンを選択してもよい。好適な技術は、ジグザグのスキャンのための 8×8 ブロックサイズを採用する。量子化された係数のジグザグのスキャンは、ゼロ値の大きいランに遭遇する機会を改良する。このゼロのランは、本質的に減少する確率を有し、ハフマン符号を使用して、効率的に符号化してもよい。

【0062】

順番に並べられ、量子化された A C 係数のストリームは、可変長コーダーに供給される。A C 成分値は、ハフマン符号またはゴロンブーライズ符号化のいずれかを用いて符号化してもよい。D C 成分値の場合、ゴロンブーライズ符号化が利用される。ランレングスコーダーは、係数をゼロの係数と非ゼロの係数に分けるが、詳細は図 6 に対して記載される。一実施形態において、ゴロンブーライズ符号化が利用される。ゴロンブーライズ符号化は指数分布を有する負でない整数を符号化する際に効率的である。ゴロンブ符号を用いることにより、指数的に分布される変数に対してより短いレングスコードを供給する際の圧縮のためにより最適である。

【0063】

ゴロンブ符号化ランレングスにおいて、ゴロンブ符号は負でない整数 m によりパラメータ化される。例えば、パラメータ m を与えられると、正の整数 n のゴロンブ符号化は、単項式符号の n/m の商により表されその後に変更されたバイナリコードにより表される剰余が続く。この剰余は、剰余が

【数 5】

$$2^{\lceil \log_2 m \rceil} - m$$

【0064】

以下ならば、

10

20

30

40

50

【数 6】

$$\lfloor \log_2 m \rfloor$$

【0065】

ビット長であり、そうでなければ

【数 7】

$$\lfloor \log_2 m \rfloor$$

10

【0066】

ビット長である。ゴロンブーライズ符号化は、パラメータ m が $m = 2^k$ として表される場合のゴロンブ符号化の特別な場合である。そのような場合、 n/m の商は整数 n のバイナリ表示を k ビットだけ右にシフトすることにより得られ、 n/m の残りは n の最小の k ビットにより表される。従って、ゴロンブーライズ符号は 2 つの連結である。ゴロンブーライズ符号化は、

【数 8】

$$p_{\alpha}(x) = c\alpha^{|x|} \quad (1)$$

20

【0067】

により与えられる二辺の幾何学的な（指数関数）分布を有した正および負の整数を符号化するために使用することができる。

【0068】

(1) において、 α は x の確率の減少を特徴づけるパラメータであり、 c は正規化定数である。 $p_{\alpha}(x)$ は単調であるので、整数値の系列は、

【数 9】

$$p_{\alpha}(x_i = 0) \geq p_{\alpha}(x_i = -1) \geq p_{\alpha}(x_i = +1) \geq p_{\alpha}(x_i = -2) \geq \dots \quad (2)$$

30

【0069】

を満足することがわかる。

【0070】

図 4 A、4 B、4 C および 5 A、5 B、5 C に図解するように、量子化された係数マトリクスにおけるゼロランおよび振幅は、指数分布を有する。これらの図に図解される分布は、実像からのデータに基づく。図 4 A は、相対周波数に対するゼロランレングスの Y 成分分布 4 0 0 を図解する。同様に、図 4 B および 4 C は、それぞれ相対周波数 4 1 0 および 4 2 0 に対するゼロランレングスの C b および C r 成分分布を図解する。図 5 A は、相対周波数に対する振幅サイズの Y 成分分布 5 0 0 を図解する。同様に、図 5 b および 5 c はそれぞれ、相対周波数に対する振幅サイズの C b 成分分布および C r 成分分布を図解する。図 5 A、5 B、および 5 C において、プロットは、DCT 係数のサイズの分布を表すことに留意する必要がある。各サイズは、係数値のレンジを表す。例えば、4 のサイズ値は、レンジ $\{-15, -14, \dots, -8, 8, \dots, 14, 15\}$ 、合計 16 の値を有する。同様に、10 のサイズ値は、レンジ $\{-1023, -1022, \dots, -512, 512, \dots, 1022, 1023\}$ 、合計 1024 の値を有する。ランレングスと振幅サイズの両方が指数分布を有することは図 4 A、4 B、4 C、5 A、5 B および 5 C からわかる。振幅の実際の分布は、以下の式 (3) に適合するように示すことができ

40

50

る：

【数 1 0】

$$p(X_{k,1}) = \frac{\sqrt{2\lambda}}{2} \exp\left\{-\sqrt{2\lambda}|X_{k,1}|\right\}, k, 1 \neq 0 \quad (3)$$

【0 0 7 1】

(3)において、 $X_{k,1}$ はそれぞれ、垂直次元および水平次元における周波数 k および1に対応するDCT係数を表し、平均値 μ_x は

【数 1 1】

10

$$\mu_x = \frac{1}{\sqrt{2\lambda}}$$

【0 0 7 2】

を表し、分散

【数 1 2】

$$\sigma^2_x$$

20

【0 0 7 3】

は、

【数 1 3】

$$\sigma^2_x = \frac{1}{2\lambda}$$

【0 0 7 4】

を表す。したがって、記載した方法で、ゴロンブライズ符号化を使用することはDCT 30
sにおいてデータを処理する際に、より最適である。

【0 0 7 5】

以下は、画像データの圧縮に対して記載されるけれども、実施形態は、等価的にオーディオデータを圧縮する実施形態に適用可能である。画像データを圧縮する際に、例えば、画像信号またはビデオ信号は、線形のまたはログ符号化されたピクセル値を有したRGB成分、またはYIQ成分またはYCbCr成分であってもよい。

【0 0 7 6】

図6はゼロ係数および非ゼロ係数のプロセス600を図解する。DCTマトリクスがスキャンされると、ゼロ係数および非ゼロ係数が別個に処理され、分離される604。ゼロデータの場合、ゼロランの長さが決定される608。ラン長は正の整数であることに留意する必要がある。例えば、ラン長が n であるとわかると、ゴロンブパラメータ m が決定される612。実施形態において、ゴロンブパラメータは、ラン長スの関数として決定される。他の実施の形態において、ゴロンブパラメータ(m)は以下の式(4)により決定される。

【数 1 4】

$$m = \lceil \log_2 n \rceil \quad (4)$$

50

【0077】

状況に応じて、ランレングスの長さおよび関連するゴロンブパラメータはカウンタまたはレジスタによりカウントされる616。nのラベルがつけられたゼロのランレングスを符号化するために、商が符号化される620。一実施形態において、商はゼロのランレングスの関数およびゴロンブパラメータとして決定される。他の実施の形態において、商(Q)は、以下の式(5)により決定される。

【数15】

$$Q = \lfloor n / 2^m \rfloor \quad (5)$$

10

【0078】

一実施形態において、商Qは実施例において、商Qは単項コードで符号化される。これはQ+1ビットを必要とする。次に、剰余が符号化される624。一実施形態において、剰余はランレングスおよび商の関数として符号化される。他の実施の形態において、剰余(R)は、以下の式(6)を用いて決定される。

【数16】

$$R = n - 2^m Q \quad (6)$$

20

【0079】

他の実施形態において、剰余Rはmビットバイナリコードで符号化される。その後、商Qおよび剰余Rが決定され、QおよびRに対する符号が連結される628、nのラベルが付けられたランレングスのための全体のコードを表す。

【0080】

非ゼロ係数もゴロンブライスをを用いて符号化される。係数振幅は正または負でありえるので、サインビットを用いる必要があり、与えられた振幅の絶対値を符号化する必要がある。非ゼロ係数の振幅がxであると仮定すると、振幅の絶対値およびサインの関数として表しても良い。したがって、振幅は以下の式(7)を用いてyとして表すことができる。

30

【数17】

$$y = \begin{cases} 2x, & \text{if } x \geq 0 \\ 2|x| - 1, & \text{otherwise} \end{cases} \quad (7)$$

【0081】

従って、非ゼロ係数の値はカウンタまたはレジスタにより最適にカウントされる632。次に、振幅がゼロより大きいまたは等しいかが決定される。そうであるなら、与えられた値の2倍の値として符号化される640。そうでない場合には、絶対値の2倍の値より1少ない値として符号化される644。他のマッピングスキームを使用してもよいことが予想される。キーポイントは、値の符号を区別する余分なビットが必要ないということである。

40

【0082】

式(7)により表される振幅を符号化することにより、xの正の値が偶数の整数となり、負の値が奇数の整数となる。さらに、このマッピングは(2)のxの確率割当てを保存する。式(7)で図示したように、符号化の利点は、正の数および負の数を表すために符号ビットを用いて回避することを実現する。マッピングが成された後、yはゼロランに対して成されたと同様の方法で符号化される。全ての係数が現在のブロックにおいて走査されるまで、手続は続けられる。

50

【0083】

この発明の実施形態は、式(1)乃至(7)の関数として係数およびランレングスの値を決定することであるけれども、式(1)

乃至(7)そのものを使用する必要はない。画像データおよびオーディオデータのより効率的な圧縮を可能にすることは、ゴロンブライスコードおよびDCT係数の指数分布の利用である。

【0084】

符号化の後のゼロランは、ゼロラン振幅と区別できないので、固定長の特別な前につけるコードを用いて最初のゼロランの発生をマークすることが必要かもしれない。振幅が非ゼロになった後では、ブロック内がすべてゼロになることはよくあることである。そのような場合、ゴロンブライスコードよりもエンドオブブロック(EOB)コードと呼ばれるコードを使用することはより効率的かもしれない。EOBコードは、また一方、情況に応じて特定の固定長コードである。

【0085】

式(1)または(3)によれば、DCT係数マトリクス内の振幅またはランレングスの確率分布は α 又は λ によりパラメータ化される。言外の意味は、特定のDCT係数ブロックが生じるという情況で符号化効率が改良されるかもしれないということである。従って、関心の質を符号化するために、適当なゴロンブライスコードパラメータを使用してもよい。一実施形態において、それぞれの累積値およびそのような値が生じる対応する回数を計算するために、各ランレングスおよび振幅サイズ値に対してカウンタまたはレジスタが使用される。例えば、累積値を記憶するためのレジスタおよび蓄積されるエレメントの数が、それぞれ、 R_{r1} および N_{r1} であるなら、以下の式(6)をライズーゴロンブパラメータとして用いて、ランレングスを符号化してもよい。

【数18】

$$\left\lceil \log_2 \frac{R_{r1}}{N_{r1}} \right\rceil \quad (6)$$

【0086】

同様の手続を、振幅に対して使用してもよい。

【0087】

残差ピクセルは、最初に、ABSDCTでコードを用いて圧縮されたデータを解凍し、次に、それをオリジナルデータから減算することにより発生される。残差ダイナミックレンジが小さければ小さいほど、圧縮は高くなる。圧縮はブロックベースなので、残差もブロックベースで発生される。残差ピクセルは両面のある指数分布であって、通常ゼロに中心がある指数分布を有することでよく知られた事実である。ゴロンブライスコードはそのようなデータに対してより最適であるので、ゴロンブライスコード化手続を用いて残差データを圧縮する。しかしながら、符号化すべきランレングスが無いので、特別なコードは必要としない。さらに、EOBコードの必要がない。従って、圧縮されたデータは2つの成分から成る。一方は、損失のある圧縮器からの成分であり、他方は、損失の無い圧縮器からの成分である。

【0088】

動作系列を符号化するとき、一時的な相関性を利用する便宜も得ることができる。一時的な相関性を完全に利用するために、ピクセル置換は最初に動作により推定され、次に、残差ピクセルを得るために動作補償予測が実行される。ABSDCTは順応性のあるブロックサイズ符号化を実行するので、ブロックサイズ情報は、動作による置換の手段として代替的に使用してもよい。さらに簡単にするために、場面変更検出は使用しない。その代わり、系列内の各フレームに対して、最初にイントラフレーム圧縮されたデータが得られる。次

に、現在のフレーム DCT および以前のフレーム DCT との間の差分がブロック単位で発生される。これは、参照することにより本願明細書に組み込まれる、2001年6月7日に出願された米国特許出願シリアル番号第09/877,578にさらに詳細に記載されている。DCT領域のこれらの残差は、ハフマン符号化手続きおよびゴロンブ・ライズ符号化手続きの両方を用いて符号化される。従って、最終の圧縮された出力はフレームあたり最小のビット数を使用する圧縮された出力に対応する。

【0089】

損失の無い圧縮アルゴリズムは、損失の無い部分を取り除くことにより、再度目的を持たせ、一方のコードを他方のコードに変換するのに好適であるハイブリッドスキームである。従って、ABS DCTを用いることは、空間領域内のピクセル相関を最大化し、予測スキームにおいて使用された残差ピクセルよりも低い分散を有する残差ピクセルを生じる。全体的なシステムの損失のある部分によって、ユーザは、フレーム間処理の助けを求める必要が無くして、分布目的のために必要な品質およびデータレートを得ることができる。これはデジタル映画アプリケーションのために配布されるプログラムにおいて、特に重要である。何故なら、圧縮された部材の損失のある部分は、配布の際により高いレベルの品質を必要とするからである。

【0090】

図9はハイブリッドの損失の無い符号化装置を図解する。図10はそのような装置上で動作してもよいプロセスを図解する。オリジナルデジタル情報904は記憶装置上に常駐し、または送信される。図9の要素の多くは、図1および図2に対してより詳細に記載されている。データのフレームは、ブロックサイズ割当て要素912、DCT/DQT変換要素916および量子化器920で構成される圧縮器908に送られる。DCT/DQTがデータ上で実行された後、そのデータは周波数領域に変換される。1つの出力922において、そのデータは量子化器920により量子化され、出力924に転送される、これは、記憶装置および/またはスイッチングで構成してもよい。上述した処理のすべては、フレーム内ベースである。上記した処理の全ては、フレーム内に基づく。

【0091】

量子化器の出力はまた解凍器928に転送される。解凍器928は圧縮器のプロセスを元に戻し、BSAにより定義されたPQRデータの知識と共に、逆量子化器932およびIDQT/IDCT936を通過する。解凍器940の結果は減算器944に供給され、そこで、オリジナルと比較される。減算器944は、各ブロックに対して未圧縮ピクセルおよび圧縮されたおよび解凍されたピクセルとの間の差分とした残差画像を計算する差分器のような、種々の要素であってよい。さらに、差分器は、条件的なフレーム間符号化のために各ブロックに対してDCT領域内の残差を得るようにしてもよい。解凍されたデータとオリジナルデータとの間の比較の結果948がピクセル残差ファイルである。すなわち、結果948は、圧縮されたおよび解凍されたデータにより経験された損失を示している。従って、オリジナルデータは結果と組み合わせられた出力922に等しい。次に、結果948は順番に並べられ、ハフマンおよび/またはゴロンブ・ライズ符号器956に送られ、第2出力960として供給される。ハフマンおよび/またはゴロンブ・ライズエンコーダ956は、ゴロンブ・ライズ符号化を使用している残差画素を符号化する一種のエントロピー・エンコーダであってもよい。意思決定は、各々のフレームのための最小のビットに基づいて、フレーム内かフレーム間かを使用するべきかどうかなされる。残差のゴロンブ・ライズ符号化の使用は、システムのより高い全体的な圧縮比に導く。

【0092】

従って、損失の無いフレーム間出力は、損失のある高品質の画像ファイル(922またはA)および残りのファイル(960またはC)の組合せまたはハイブリッドである。

【0093】

フレーム間符号化も利用してもよい。量子化器の出力は、BSAの知識と共に、記憶装置964に転送される。1フレーム分のデータを集めると、減算器966は記憶したフレームを次のフレーム968と比較する。その差分はDCT残差970を生じる。この残差

10

20

30

40

50

は次に順番に並べられ、および／またはゴロンブーライズ符号化され974、第3出力データセット976を出力924に供給する。従って、BおよびCのフレーム間の損失の無いファイルは、コンパイルされる。従って、組合せ(A+CまたはB+C)のいずれかをサイズの考察に基づいて選択してもよい。さらに、純粋なフレーム内出力は、編集目的のために望ましいかもしれない。

【0094】

図1に戻ると、エンコーダ104により発生された圧縮された画像信号は一時的にバッファ142を用いて記憶してもよく、その次に、送信チャネル108を用いてデコーダ112に送信してもよい。送信チャネル108は磁気または光学の記憶装置のような物理的な媒体であってもよく、または、有線または無線の輸送処理または装置であってもよい。ブロックサイズ割当て情報を含むPQRデータもデコーダ112(図2)に供給される。デコーダ112はバッファ164および可変長デコーダ168を含む。可変長デコーダ168は、ランレングス値および非ゼロ値をデコードする。可変長デコーダ168は同様に動作するが、図6で記載した方法と反対の方法で動作する。

【0095】

可変長デコーダ168の出力は、採用される走査スキームに従って係数を順序づける逆シリアライザ172に供給される。例えば、ジグザグ走査、垂直走査、および水平走査の混合が使用されるなら、逆シリアライザ172は、採用された走査のタイプの知識を用いて、その係数を再順序付けするであろう。逆シリアライザ172は、PQRデータを受信し、復号係数ブロックへの係数の適切な順序付けを支援する。

【0096】

復号ブロックは、量子化器スケールファクタおよび周波数重み付けマスクの使用により処理を元に戻すために、逆量子化器174に供給される。

【0097】

微分四部木変換が適用されたなら、次に、係数ブロックはIDQTエレメント186に供給され、その後IDCTエレメント190が続く。そうでなければ、係数ブロックは、直接IDCTエレメント190に供給される。IDQTエレメントおよびIDCTエレメント190は、係数を逆変換し、ピクセルデータのブロックを生成する。従って、ピクセルデータは、補間され、RGBフォームに変換されなければならないかもしれない、そして、次に、将来の表示のために記憶される。

【0098】

図7はゴロンブーライズ符号化のための装置を図解する。図7の装置は、望ましくは、図6に対して記載されたプロセスを実施する。決定器704はランレングスおよびゴロンブパラメータ(m)を決定する。状況に応じて、各ランレングスおよび振幅サイズ値に対してカウンタまたはレジスタ708が使用され、それぞれの累積値およびそのような値が生じる対応する回数を計算する。エンコーダ712は、ランレングスおよびゴロンブパラメータの関数として商(Q)を符号化する。エンコーダ712は、また、ランレングス、ゴロンブパラメータおよび商の関数とし、剰余(R)を符号化する。他の実施形態において、エンコーダ712は、非ゼロデータ値および非ゼロデータ値の符号の関数として非ゼロデータを符号化する。連結器716を用いてQ値をR値と連結する。

【0099】

例として、種々の事例となる論理ブロック、フローチャート、および本願明細書に開示した実施形態に関連して記載したステップは、特定用途向け集積回路(ASIC)、プログラマブルロジックデバイス、ディスクリートゲートまたはトランジスタロジック、例えばレジスタおよびFIFOのようなディスクリートハードウェアコンポーネント、一組のファームウェア命令を実行するプロセッサ、いかなる一般的なプログラマブルソフトウェアおよびプロセッサ、またはそれらのいずれかの組合せを用いて実施、または実行してもよい。プロセッサは、有利にはマイクロプロセッサであってもよいが、他の形態では、プロセッサ、いかなる一般的なプロセッサ、コントローラ、マイクロコントローラまたはステートマシンであってもよい。ソフトウェアは、RAMメモリー、フラッシュメモリー、ROMメモリー

10

20

30

40

50

、レジスタ、ハードディスク、リムーバブル・ディスク、CD-ROM、DVD-ROMまたは技術的に知られている他のいずれかの形態の記憶媒体に常駐することができる。

【0100】

好適実施形態の上述の記述は、技術に熟達したいかなる人もがこの発明を製作し、または使用することができることを可能にする。これらの実施例に対する種々の変更は、当業者には容易に、明白であり、本願明細書において定義される一般的な原理は発明の能力を用いずに他の実施形態に適用してもよい。従って、本発明は本願明細書において示される実施形態に限定されることを意図したのではなく、本願明細書において開示される原理および新規な特徴に一致する最も広い範囲が許容されるべきである。

【0101】

本発明の他の特徴および利点は以下のクレームに記載される。

【図面の簡単な説明】

【0102】

【図1】図1は、画像圧縮および処理システムのエンコーダ部分のブロック図である。

【図2】図2は、画像圧縮および処理システムのデコーダ部分のブロック図である。

【図3】図3は、分散ベースのブロックサイズ割当てに關係している処理ステップを例示するフローチャートである。

【図4A】図4Aは、DCT係数マトリクスのY成分ランレングスの指数分布を図解する。

【図4B】図4Bは、DCT係数マトリクスのCb成分ランレングスの指数分布を図解する。

【図4C】図4Cは、DCT係数マトリクスのCr成分ランレングスの指数分布を図解する。

【図5A】図5Aは、Y成分の振幅サイズの指数分布またはDCT係数マトリクスにおけるY成分の振幅サイズの指数分布を図解する。

【図5B】図5Bは、Cb成分の振幅サイズの指数分布またはDCT係数マトリクスのCb成分の振幅サイズの指数分布を図解する。

【図5C】図5Cは、Cr成分の振幅サイズの指数分布またはDCT係数マトリクスにおけるCr成分の振幅サイズの指数分布を図解する。

【図6】図6は、ゴロンブライズ符号化プロセスを図解する。

【図7】図7は、ゴロンブライズ符号化のための装置を図解する。

【図8】図8は、Dc成分値を符号化するプロセスを図解する。

【図9】図9は、損失の無い圧縮のための装置を図解する。

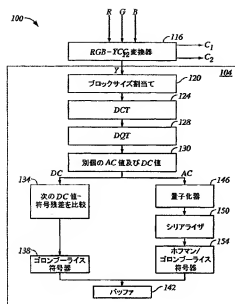
【図10】図10は、ハイブリッドの損失の無い圧縮の方法を図解する。

10

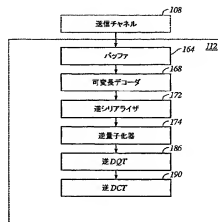
20

30

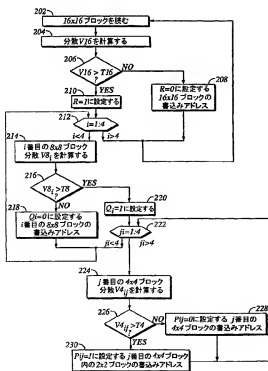
【図 1】



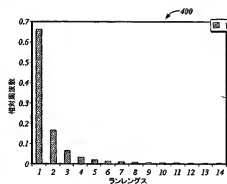
【図 2】



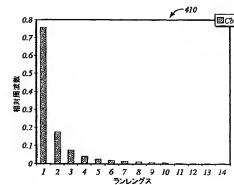
【図 3】



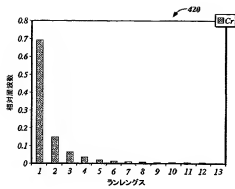
【図 4 A】



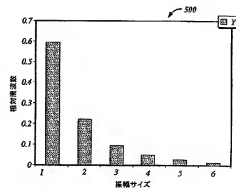
【図 4 B】



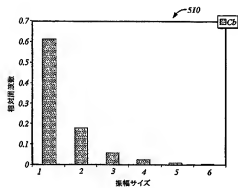
【図 4 C】



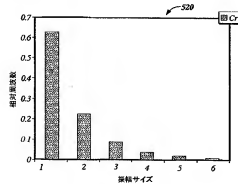
【図 5 A】



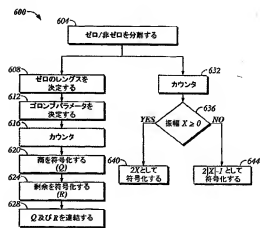
【図 5 B】



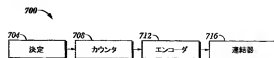
【図 5 C】



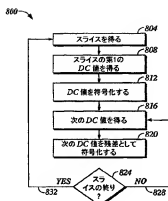
【図 6】



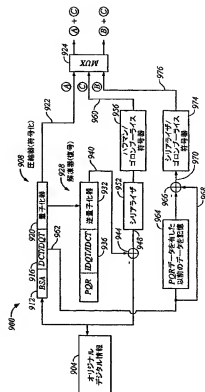
【図 7】



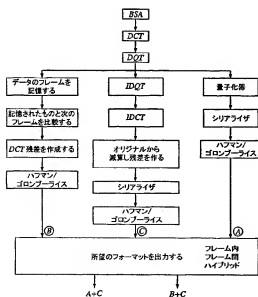
【図 8】



【図 9】



【図 10】



【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US02/21151																		
A. CLASSIFICATION OF SUBJECT MATTER IPC(7) : G06K 9/36 US CL : 382/240, 244, 236 According to International Patent Classification (IPC) or in both national classification and IPC																				
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 382/240, 244, 232, 236, 246; 375/240, 12, 240, 13, 240, 14, 240.2 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Continuation Sheet																				
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1"> <thead> <tr> <th>Category *</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>X ---</td> <td>US 4,903,317 A (NISHIHARA et al) 20 February 1990 (20.02.1990), Figs. 1-2, column 2, line 60 to column 3, line 50</td> <td>1, 2, 5, 6, 9, 10, 17, 18</td> </tr> <tr> <td>X ---</td> <td>JP 9-238345 A (SUZUKI et al) 09 September 1997 (09.09.1997), Fig. 3</td> <td>3, 4, 7, 8, 11, 12, 19 13, 14, 20, 21</td> </tr> <tr> <td>X ---</td> <td>US 5,452,104 A (LEE) 19 September 1995 (19.09.1995), column 2, line 25 to column 3, line 43</td> <td>15, 16, 22</td> </tr> <tr> <td>Y</td> <td>US 5,680,129 A (WEINBERGER et al) 21 October 1997 (21.10.1997), column 20, lines 46-67</td> <td>3, 7, 11, 15, 19, 22</td> </tr> <tr> <td>Y</td> <td></td> <td>4, 8, 12, 16</td> </tr> </tbody> </table>			Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	X ---	US 4,903,317 A (NISHIHARA et al) 20 February 1990 (20.02.1990), Figs. 1-2, column 2, line 60 to column 3, line 50	1, 2, 5, 6, 9, 10, 17, 18	X ---	JP 9-238345 A (SUZUKI et al) 09 September 1997 (09.09.1997), Fig. 3	3, 4, 7, 8, 11, 12, 19 13, 14, 20, 21	X ---	US 5,452,104 A (LEE) 19 September 1995 (19.09.1995), column 2, line 25 to column 3, line 43	15, 16, 22	Y	US 5,680,129 A (WEINBERGER et al) 21 October 1997 (21.10.1997), column 20, lines 46-67	3, 7, 11, 15, 19, 22	Y		4, 8, 12, 16
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.																		
X ---	US 4,903,317 A (NISHIHARA et al) 20 February 1990 (20.02.1990), Figs. 1-2, column 2, line 60 to column 3, line 50	1, 2, 5, 6, 9, 10, 17, 18																		
X ---	JP 9-238345 A (SUZUKI et al) 09 September 1997 (09.09.1997), Fig. 3	3, 4, 7, 8, 11, 12, 19 13, 14, 20, 21																		
X ---	US 5,452,104 A (LEE) 19 September 1995 (19.09.1995), column 2, line 25 to column 3, line 43	15, 16, 22																		
Y	US 5,680,129 A (WEINBERGER et al) 21 October 1997 (21.10.1997), column 20, lines 46-67	3, 7, 11, 15, 19, 22																		
Y		4, 8, 12, 16																		
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.																				
<table border="1"> <thead> <tr> <th>* Special categories of cited documents:</th> <th>T</th> <th>Y</th> </tr> </thead> <tbody> <tr> <td>"A" document defining the general state of the art which is not considered to be of particular relevance</td> <td>"T" later documents published after the international filing date or priority date and not in conflict with the application but cited to understand the principles or theory underlying the invention</td> <td></td> </tr> <tr> <td>"B" earlier applications or patents published on or after the international filing date</td> <td>"X" document of particular relevance; the claimed invention cannot be considered novel, or cannot be considered to involve an inventive step when the document is taken alone</td> <td></td> </tr> <tr> <td>"C" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another claim(s) or other special reason (as specified)</td> <td>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</td> <td></td> </tr> <tr> <td>"O" document referring to an oral disclosure, use, exhibition or other means</td> <td>"A" document member of the same patent family</td> <td></td> </tr> <tr> <td>"P" document published prior to the international filing date but later than the priority date claimed</td> <td></td> <td></td> </tr> </tbody> </table>			* Special categories of cited documents:	T	Y	"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later documents published after the international filing date or priority date and not in conflict with the application but cited to understand the principles or theory underlying the invention		"B" earlier applications or patents published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel, or cannot be considered to involve an inventive step when the document is taken alone		"C" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another claim(s) or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art		"O" document referring to an oral disclosure, use, exhibition or other means	"A" document member of the same patent family		"P" document published prior to the international filing date but later than the priority date claimed		
* Special categories of cited documents:	T	Y																		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later documents published after the international filing date or priority date and not in conflict with the application but cited to understand the principles or theory underlying the invention																			
"B" earlier applications or patents published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel, or cannot be considered to involve an inventive step when the document is taken alone																			
"C" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another claim(s) or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art																			
"O" document referring to an oral disclosure, use, exhibition or other means	"A" document member of the same patent family																			
"P" document published prior to the international filing date but later than the priority date claimed																				
Date of the actual completion of the international search 12 September 2002 (12.09.2002)		Date of mailing of the international search report 26 DEC 2002																		
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20531 Facsimile No. (703) 305-3230		Authorized officer Wengping Chen Telephone No. 703 306-2796																		

INTERNATIONAL SEARCH REPORT

PCT/US02/21151

Continuation of B. FIELDS SEARCHED Item 3:
EAST(USPAT;US-PGPUB,EPO,JPO,DERWENT,IBM,TDB,USOCR): DQT, discrete quadtree transform, error image, residual,
leaky, lossless

フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KC,KZ,MD,RU,TJ,TM),EP(AT, BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW, ML,MR,NE,SN,TD,TC),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES, FI,GB,GD,GE,GH,GM,HR,HU,IE,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,ME,MK,MN,MW,MY,MZ,N O,NZ,OM,PH,PL,PT,RO,RU,SE,SG,SI,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,UZ,VN,YU,ZA,ZM,ZW

(74)代理人 100084618

弁理士 村松 貞男

(74)代理人 100092196

弁理士 橋本 良郎

(72)発明者 ラビーン・ドラ、ビジャヤラクシュミ・アール

アメリカ合衆国、カリフォルニア州 92130、サン・ディエゴ、カレ・マー・デ・パレナス 4272

(72)発明者 スヤガラジャン、カダヤム

アメリカ合衆国、カリフォルニア州 92130、サン・ディエゴ、サーストン・プレイス 482

(72)発明者 ラツツェル、ジョン

アメリカ合衆国、カリフォルニア州 92129、サン・ディエゴ、ビムバーネル・ウェイ 12823

(72)発明者 モーレイ、ステイブン・エー

アメリカ合衆国、カリフォルニア州 92025、エスコンディッド、ロスト・オーク・レーン 540

Fターム(参考) 5C059 MA05 MA23 MA45 MC11 ME05 PP04 PP15 PP16 UA02

5J064 AA02 BA09 BA16 BC02 BC04 BC05 BC08 BC09 BC14 BC16

BC21 BC22 BC29 BC30 BD02

【公報種別】特許法第17条の2の規定による補正の掲載
 【部門区分】第7部門第3区分
 【発行日】平成18年1月5日(2006.1.5)

【公表番号】特表2005-520360(P2005-520360A)
 【公表日】平成17年7月7日(2005.7.7)
 【年通号数】公開・登録公報2005-026
 【出願番号】特願2003-511463(P2003-511463)
 【国際特許分類】

H 0 4 N 7/32 (2006.01)
 H 0 3 M 7/30 (2006.01)
 H 0 4 N 7/30 (2006.01)

【F I】

H 0 4 N 7/137 Z
 H 0 3 M 7/30 A
 H 0 4 N 7/133 Z

【誤訳訂正書】
 【提出日】平成17年8月2日(2005.8.2)
 【誤訳訂正1】

【訂正対象書類名】特許請求の範囲
 【訂正対象項目名】全文
 【訂正方法】変更
 【訂正の内容】
 【特許請求の範囲】

【請求項1】

下記を具備する、画像を表す信号を損失無く圧縮し、符号化する方法：

損失のある圧縮されたデータファイルが発生する；

残差の圧縮されたデータファイルが発生する；および

損失のあるデータファイルを前記残差データファイルと結合し、損失の無いデータファイルを作成する、前記損失の無いデータファイルは、実質的にオリジナルデータファイルと同一である。

【請求項2】

前記損失のある圧縮されたデータファイルおよび前記残差の圧縮されたデータファイルはフレーム内に基づいてまたはフレーム間に基づいて発生される、請求項1に記載の方法。

【請求項3】

発生は、ディスクリットコサイン変換(DCT)技術およびディスクリット四部木変換(DQT)技術を利用する、請求項1に記載の方法。

【請求項4】

発生は、ゴロンプライズ符号化技術を利用する、請求項1に記載の方法。

【請求項5】

下記を具備する、画像を損失無く圧縮し符号化する装置：

損失のある圧縮されたデータを発生する手段；

残差の圧縮されたデータファイルが発生する手段；および

前記損失のあるデータファイルと、前記残差データファイルを結合し、損失の無いデータファイルを作成する手段、前記損失の無いデータファイルは実質的にオリジナルデータファイルと同一である。

【請求項6】

前記損失のある圧縮されたデータファイルが発生する手段および前記残差の圧縮された

データファイルを発生する手段は、フレーム内に基づいて、またはフレーム間に基づいて発生される、請求項 5 に記載の装置。

【請求項 7】

前記発生手段は、ディスクリットコサイン変換（DCT）技術およびディスクリット四部木変換（DQT）技術を利用する、請求項 5 に記載の装置。

【請求項 8】

前記発生手段は、ゴロンブライズ技術を利用する、請求項 5 に記載の方法。

【請求項 9】

下記を具備する、画像を表す信号を損失無く、圧縮し、符号化する方法：

画像を表す信号を圧縮し、それにより、画像の圧縮されたバージョンを作成する；

前記画像の量子化された圧縮されたバージョンを順番に並べ、それにより、画像の順番に並べられた量子化された圧縮されたバージョンを作成する；

前記画像の圧縮されたバージョンを解凍する；

前記画像と前記画像の解凍されたバージョンとの間の差分を決定し、それにより、前記画像の残差バージョンを作成し；および

前記画像の損失のあるバージョンと前記画像の残差バージョンを出力し、前記画像の損失のあるバージョンと前記画像の残差バージョンとの組合せが実質的にオリジナル画像と同じである。

【請求項 10】

前記損失の無い圧縮はフレーム間に基づく、請求項 9 記載の方法。

【請求項 11】

圧縮は、ディスクリットコサイン変換（DCT）技術およびディスクリット四部木（DQT）変換技術を利用する、請求項 9 記載の方法。

【請求項 12】

順番に並べることはゴロンブライズ符号化技術を利用する、請求項 9 記載の方法。

【請求項 13】

下記を具備する、複数のフレームから構成される画像情報を表す信号を損失無く圧縮し、符号化する方法：

第 1 フレームを圧縮し、それにより、前記画像の圧縮されたバージョンを作成する；

前記画像の圧縮されたバージョンを量子化し、それにより、前記画像の損失のあるバージョンを作成する；

前記画像の量子化された圧縮されたバージョンを順番に並べ、それにより、前記画像の順番に並べられた量子化された圧縮されたバージョンを作成する；

前記画像を表す信号の第 2 フレームを圧縮する；

前記画像の前記第 1 フレームと前記第 2 フレームとの間の差分を決定し、それにより、前記画像の残差バージョンを作成する；および

前記画像の残差バージョンを有した前記画像の損失のあるバージョンを出力する、前記画像の損失のあるバージョンと前記画像の残差バージョンとの組合せは実質的にオリジナル画像と同じである。

【請求項 14】

前記損失の無い圧縮はフレーム間に基づく、請求項 13 に記載の方法。

【請求項 15】

圧縮はディスクリットコサイン変換（DCT）技術およびディスクリット四部木変換（DQT）技術を利用する、請求項 13 に記載の方法。

【請求項 16】

順番に並べることはゴロンブライズ符号化技術を利用する、請求項 13 に記載の方法。

【請求項 17】

画像を表す信号を損失無く圧縮し、符号化する方法：

画像を表す信号を圧縮し、それにより、画像の圧縮されたバージョンを作成する手段；

前記画像の圧縮されたバージョンを量子化し、それにより前記画像の損失のあるバージョンを作成する手段；

前記画像の前記量子化された圧縮されたバージョンを順番に並べ、それにより前記画像の順番に並べられた量子化された圧縮されたバージョンを作成する手段；

前記画像の圧縮されたバージョンを解凍する手段；

前記画像と前記画像の解凍されたバージョンとの間の差分を決定し、それにより、前記画像の残差バージョンを作成する手段；および

前記画像の損失のあるバージョンと前記画像の残差バージョンとを出力する手段、前記画像の損失のあるバージョンと前記画像の残差バージョンの組合せは実質的にオリジナル画像と同じである。

【請求項18】

前記損失の無い圧縮はフレーム間に基づく、請求項17記載の装置。

【請求項19】

前記圧縮手段はディスクリートコサイン変換(DCT)技術およびディスクリート四部木変換(DQT)技術を利用する、請求項17記載の装置。

【請求項20】

20. 下記を具備する、画像情報を表す信号を損失無く圧縮し、符号化する装置：

第1フレームを圧縮し、それにより前記画像の圧縮バージョンを作成する手段；

前記画像の前記圧縮されたバージョンを量子化し、それにより前記画像の損失のあるバージョンを作成する手段；

前記画像の前記量子化された圧縮されたバージョンを順番に並べ、それにより、前記画像の順番に並べられた量子化された圧縮されたバージョンを作成する手段；

前記画像を表す信号の第2フレームを圧縮する手段；

前記画像の前記第1フレームと前記第2フレームとの間の差分を決定し、それにより、前記画像の残差バージョンを作成する手段；および

前記画像の残差バージョンを有した前記画像の損失のあるバージョンを出力する手段、前記画像の損失のあるバージョンと前記画像の残差バージョンとの組合せは実質的にオリジナル画像と同じである。

【請求項21】

前記損失の無い圧縮は、フレーム間に基づく、請求項20に記載の装置。

【請求項22】

前記圧縮する手段はディスクリートコサイン変換(DCT)技術およびディスクリート四部木変換(DQT)技術の組合せを利用する、請求項20に記載の装置。

【誤訳訂正2】

【訂正対象書類名】明細書

【訂正対象項目名】全文

【訂正方法】変更

【訂正の内容】

【書類名】明細書

【発明の名称】損失の無い方法でデジタル画像を符号化するための装置および方法

【技術分野】

【0001】

本発明は、画像処理および圧縮に関する。特に、本発明は周波数領域のビデオ画像および音声情報の損失の無い符号化に関する。

【背景技術】

【0002】

デジタル画像処理は、デジタル信号処理の一般的な専門分野において、重要な位置を有する。人間の視覚認知の重要性は、デジタル画像処理の技術および科学において、著しい関心および進展を促進させた。フィルムまたは映画を投影するために使用されるビデオ信号のようなビデオ信号の送受信の分野において、種々の改良が画像圧縮技術にな

されつつある。現在のおよび提案されたビデオシステムの多くは、デジタル符号化技術を利用する。この分野の視点は、画像符号化、画像復元および画像特徴選択を含む。画像符号化は、効率的な方法でデジタル通信の画像を送信するための試みを表し、同時に、必要とされる帯域幅をできるだけ最小にするように数ビットを使用し、ある制限内で歪みを維持する。画像復元は物体の真の画像を再生するための努力をあらわす。通信チャンネル上に送信される符号化された画像は種々の要因により歪みを生じるかもしれない。質が低下する原因は、もともと物体から画像を創出する際に生じていたかもしれない。特徴選択は、画像のある属性の選択に言及する。そのような属性は、より広い文脈において、認識、分類および決定において必要かもしれない。

【0003】

デジタル映画のようなビデオのデジタル符号化は、改良された画像圧縮技術から利益を得る領域である。デジタル画像圧縮は、一般的に、2つのカテゴリに分類されるかもしれない。すなわち、損失の無い方法と損失のある方法である。損失の無い画像は、いかなる情報の損失も無しに再生される。損失のある方法は、以下を含む：圧縮比、圧縮アルゴリズムの質、およびアルゴリズムの実施に応じて多少の情報の回復できない損失。一般に、費用効率が低いデジタル映画手法のために要望される圧縮比を得るために損失のある圧縮手法が考慮される。デジタル映画の品質レベルを得るために、圧縮手法は、視覚的に損失の無いレベルの性能を提供しなければならない。従って、圧縮処理の結果として情報の数学的な損失があるけれども、この損失によって生じた画像歪みは通常の画像観察下で視聴者に感知できないようであればならない。

【0004】

既存のデジタル画像圧縮技術は、ほかの応用のために開発されてきた、すなわち、テレビシステムのために開発されてきた。そのような技術は、意図したアプリケーションに対して設計上の妥協を適当なものとしたが映画プレゼンテーションに必要な品質要件に適さない。

【0005】

デジタル映画圧縮技術は、映画ファンが以前経験した視覚の質を提供しなければならない。理想的には、デジタル映画の視覚品質は、高品質の公開版フィルムの品質を超えようと試みなければならない。同時に、この圧縮技術は高符号化効率を実用的にさせなければならない。本願明細書において定義されるように、コーディング効率はある質的レベルを満たすために圧縮された画像品質に必要なビットレートに言及する。さらに、システムおよび符号化技術は、異なるフォーマットに対応するように内蔵の柔軟性を有さなければならない。すなわち、小さなサイズで効率的なデコーダまたはエンコーダ処理でなければならない。

【0006】

入手可能な多くの圧縮技術は、有意水準の圧縮を提供するが、結果として、ビデオ信号の品質の劣化を生じる。典型的に、圧縮された情報を転送するための技術は、圧縮された情報が一定のビットレートで転送されることを必要とする。

【0007】

ビデオ信号に対して所望のレベルの品質を維持しながら、有意水準の圧縮を提供することができる1つの圧縮技術は、符号化された離散コサイン変換(DCT)係数データの順応して大きさに合わせて作られたブロックおよびサブブロックを利用する。この技術は、以下において、順応ブロックサイズ離散コサイン変換(ABSDCT)方法と呼ばれるであろう。この技術は、この発明の譲受人に譲渡され、参照することにより、本願明細書に組み込まれる米国特許番号第5,021,891号、(発明の名称:「順応ブロックサイズ画像圧縮方法およびシステム」)に開示される。DCT技術は、また、この発明の譲受人に譲渡され、参照することにより本願明細書に組み込まれる米国特許番号第5,107,345号(発明の名称:「順応ブロックサイズ圧縮方法およびシステム」)にも開示される。さらに、微分四部木と組み合わせたABSDCT技術も、この発明の譲受人に譲渡され、参照することにより、本願明細書に組み込まれる、米国特許番号第5,452,104号(発明の名称:「順応

ブロックサイズ画像圧縮方法およびシステム」)に開示されている。これらの特許において開示されるシステムは、「フレーム内」符号化と呼ばれるものを利用し、画像データの各フレームはその他のフレームの内容に関係無く符号化される。ABSDCT技術を用いて、各1送けられるデータレートは、画質の識別可能な劣化のない約15億ビット/秒からおおよそ5000万ビット/秒まで、低減してもよい。

【0008】

ABSDCT技術は、白黒カラー画像を圧縮するかまたは画像を表す信号を圧縮するために使用してもよい。カラー入力信号は、YIQフォーマットであってもよく、Yは輝度または光度サンプルであり、IおよびQは、各4:4:4または代替フォーマットに対して、クロミナンスまたはカラーサンプルである。YUV、YCbCrまたはRGBフォーマットのような他の周知のフォーマットも、使用してもよい。色に対する目の低空間の感度のため、大部分の調査は、色成分の副サンプルが水平および垂直方向に4だけ合理的であることを示した。したがって、ビデオ信号は4つの輝度サンプルおよび2つのクロミナンス・サンプルによって表されることができる。

【0009】

ABSDCTを用いて、ビデオ信号は一般に処理のため、ピクセルのブロックに分割されるであろう。各々のブロックに対して、輝度およびクロミナンス成分は、ブロックサイズ割当てエレメントまたはブロックインターリーバーに渡される。例えば、 16×16 (ピクセル) ブロックは、ブロックインターリーバーに渡してもよい。ブロックインターリーバーは各 16×16 ブロック内で画像サンプルを順序付けるかまたは組織化して離散コサイン変換(DCT)解析のためにデータのブロックおよび合成副ブロックを生成する。DCTオペレータは、時間および空間のサンプル信号を同じ信号の周波数表示に変換する1つの方法である。周波数表示に変換することによって、DCT技術は、量子化器が画像の周波数分布特性を利用するように設計することができるように、非常の高レベルの圧縮を可能にすることを示してきた。好適実施形態において、1つの 16×16 DCTは第1の順序づけに適用され、4つの 8×8 DCTは第2の順序づけに適用され、 16 の 4×4 DCTは、第3の順序づけに適用され、および第4の 2×2 DCTは第4の順序づけに適用される。

【0010】

DCT操作は、ビデオソースに固有の空間冗長度を低減する。DCTが実行された後、殆どのビデオ信号エネルギーは、2乃至3のDCT係数に集中する傾向がある。さらなる変換、微分四部木変換(DQT)は、DCT係数の中で冗長度を低減するために使用してもよい。

【0011】

16×16 のブロックおよび各副ブロックに対して、DCT係数値および(DQTが使用されるなら)DQT値が解析され、ブロックまたは副ブロックを符号化するのに必要なビット数を決定する。次に、符号化するために最少数のビットを必要とする副ブロックの組合せが選択され、画像セグメントを表す。例えば、2つの 8×8 副ブロック、6つの 4×4 副ブロックおよび8つの 2×2 副ブロックは、画像セグメントを表すために選択してもよい。

【0012】

次に、選択されたブロックまたは副ブロックの組合せは適切に順番に 16×16 のブロックに配列される。次に、DCT/DQT係数値は、送信の準備のために、周波数重みづけされ、量子化され、(可変長符号化)のような符号化をしてもよい。上記したABSDCT技術は、著しくよく実行するけれども、それはコンピュータを必要とするほど強力である。

【0013】

さらに、上述したABSDCTの使用は視覚的に損失が無いけれども、時には、データが符号化される正確な方法でデータを再生することが望ましい。例えば、マスタリングおよび永久保存の目的は、正確にもとの状態に再生することができるようにデータを圧縮する必要がある。

【0014】

典型的に損失の無い画像のための圧縮システムは、予測子から構成され、予測子は符号化される現在のピクセルの値を推定する。残りのピクセルは実際のピクセルと予測されるピクセルとの間の差分として得られる。次に、残りのピクセルはエントロピー符号化され記憶されまたは送信される。予測はピクセルの相関を除去するもので、残りのピクセルは特徴のある両面のある指数（ラブラシアン）分布を有した低減されたダイナミックレンジを有する。それゆえに、圧縮を有する。残りの圧縮量は、予測および次のエントロピー符号化方法に依存する。大部分の共通して使う予測方法は、差分PCM（DPCM）および適応DPCM（ADPCM）のようなその変形である。

【発明の開示】

【発明が解決しようとする課題】

【0015】

ベルに基づく予測を伴う問題は、残差がまだ高エネルギーを有するという点である。それは、少数の隣接したピクセルだけが予測方法において使われるという事実による。したがって、ベルに基づく予測スキームの符号化効率を改良する余地がある。

【課題を解決するための手段】

【0016】

本発明の実施例は、圧縮を達成するために、損失の無い方法でデジタル画像およびビデオデータを圧縮するためのシステムを記載する。システムは、ハイブリッドである。このことは、損失のある方法で前記データを圧縮する一部と損失の無い態様で残差データを圧縮する一部とを有することを意味する。損失のある部分に対して、システムは適応ブロックサイズ離散コサイン変換（ABSDCT）アルゴリズムを使用する。ABSDCTシステムは、高い表示品位および圧縮比を生じる前記データを圧縮する。残像は、オリジナルと、ABSDCTシステムから解凍されたものとの間の差分として得られる。この残差は、ゴロンブライス（Golomb-Rice）符号化アルゴリズムを用いて損失無く符号化される。視覚的に基づいた適応ブロックサイズおよびDCT係数の量子化により、残差は非常に低いエネルギーを有し、従って、全体に損失の無い圧縮比を生じる。

【0017】

ABSDCTシステムは、映画の品質において、高い圧縮比を得る。ABSDCTシステムはブロックに基づくので、ABSDCTシステムはいかなるベルに基づくスキームよりも非常によいピクセル相関を除去する。したがって、ABSDCTシステムは本願明細書において記載される損失の無いシステムにおける予測として使用される。この予測子と連動して、損失の無い符号化システムが加えられ、ハイブリッドの損失の無い圧縮システムを形成する。留意すべきは、システムは、動画像同様に静止画像を圧縮することができるという点である。静止画像なら、ABSDCT圧縮データおよびエントロピー符号化された残差データのみが圧縮された出力として使用される。運動順序のために、フレーム内圧縮を使用するかまたはフレーム間圧縮を使用するかは決定がなされる。例えば、 $f(t)$ が時刻 t における画像フレームを表すなら、 $F(t)$ および $F(t + \Delta t)$ は、それぞれ時刻 t および $t + \Delta t$ における画像フレームのDCTを示す。 Δt が2つの連続的なフレーム間の時間間隔と対応する点に注意する必要がある。

【0018】

この発明は、データが符号化された正確な方法でデータを再生することを可能にするデータを圧縮するための装置および方法に具現化される。実施形態は、フレーム内符号化、フレーム間符号化または、それら2つのハイブリッドを実行するシステムを含む。このシステムは、離散コサイン変換係数データの順応した大きさに合わせて作られたブロックおよび副ブロックを利用する品質に基づいたシステムである。1つのブロックのピクセルデータは、エンコーダに対する入力である。エンコーダは、ブロックサイズ割当て（BSA）構成要素を含み、ブロック割当て構成要素は処理のためにピクセルの入力ブロックを分割する。ブロックサイズ割当ては、入力ブロックおよびさらに再分割されたブロックの分散に基づく。一般に、ブロックおよび副ブロックの平均値が異なる所定の範囲に入るなら

、より大きい分散を伴うエリアは、より小さなブロックに再分割され、より小さな分散を伴うエリアは再分割されない。従って、最初に、ブロックの分散閾値は、その平均値に応じて公称値から変更され、次に、ブロックの分散はこの閾値と比較され、そして、分散が閾値より大きい場合、そのブロックは再分割される。

【0019】

ブロックサイズ割当ては変換構成要素に供給され、変換構成要素はピクセルデータを周波数領域データに変換する。この変換は、ブロックサイズ割当てを介して選択されたブロックおよび副ブロックのみに実行される。A C 構成要素の場合、量子化および連続化を介してその変換データにスケーリング処理をする。変換データの量子化は、コントラスト、係数カウント、レート歪み、ブロックサイズ割当ての密度及び／または過去スケールファクタに対して調節するスケールファクタのような、画像品質測定基準に基づいて量子化される。ジグザグスキャンのような連続化は同じ値の最も長い可能なランレングスを作ることに基づく。次に、データのストリームは送信に備えて、可変長符号化器により符号化される。コーディングはハフマン・コーディングであってもよい、または、ゴロンブーライズ符号化のような指数分布に基づいてもよい。

【0020】

A B S D C T のようなハイブリッド圧縮システムの使用は、ピクセルまたは D C T 値の良好な予測子のように動作する。それゆえ、A B S D C T は、ペルに基づく予測を使用しているシステムより高い損失の無い圧縮比を生じる。損失のある部分は、ディジタル映画品質結果を供給する。すなわち、視覚的に損失の無いファイルに圧縮結果を供給する。損失の無い部分の場合、ハフマンコードと異なり、ゴロンブーライズ符号化はいかなる先験的符号発生を必要としない。それゆえ、ゴロンブーライズ符号化は、ハフマン符号化の場合に記憶される拡張コードブックを必要としない。これは、チップの有効スペースの効率的な使用を生じる。それゆえに、チップ・サイズは、ハードウェア実施において低減される。さらに、ゴロンブーライズ符号化は、ハフマン符号化より実施するのにより簡単である。また、D C T 係数または残りのピクセルは必然的に指数分布を有するので、ゴロンブーライズ符号化は、ハフマン符号化よりも高い符号化効率を得る。さらに、圧縮システムの損失のある部分がブロック再分割において視覚的に重要な情報を使用するので、コンテキストモデリングは残差符号化において固有である。これは、残差符号化のためにコンテキストデータを集めるのに、余分な記憶レジスタが必要ないという点で、重要である。また、運動推定が使われないので、システムは実施するのに非常に簡単である。

【0021】

画像情報を表す信号を損失が無いように圧縮し符号化するための装置および方法が請求される。画像情報を表す信号は、圧縮され、画像の圧縮されたバージョンを作る。画像の圧縮されたバージョンは量子化され、それにより、画像の損失のあるバージョンを作る。画像の圧縮されたバージョンは、また、順番に並べられ、画像の順番に並べられた量子化された圧縮されたバージョンを作る。このバージョンは、次に、解凍され、オリジナル画像と解凍されたバージョンとの間の差分が決定され、それにより画像の残差バージョンを作る。画像の損失のあるバージョンおよび画像の残差バージョンは、別個にまたは結合されて別に出力してもよく、その場合、画像の解凍された損失のあるバージョンと、画像の残差バージョンとの組合せは実質的にオリジナル画像と同じである。

【0022】

画像情報を表す信号を損失無く圧縮し、符号化する方法が、請求される。損失のある圧縮されたデータファイルおよび残りの圧縮されたデータファイルが発生される。損失のある圧縮されたデータファイルおよび残りの圧縮されたデータファイルが結合されると、実質的にオリジナルデータファイルと同一であり損失の無いデータファイルが作られる。

【0023】

したがって、効率的に損失の無い圧縮を提供するための装置および方法を提供することが、実施形態の1つの観点である。

【0024】

マスタリングおよびアーカイブの目的の助けとなる方法で、損失無く、デジタル画像情報およびオーディオ情報を圧縮することは実施形態の他の観点である。

【0025】

フレーム間に基づいて、損失の無い圧縮システムを提供することは実施形態の他の観点である。

【0026】

フレーム内に基づいて損失の無い圧縮システムを提供することは実施形態の他の観点である。

【発明を実施するための最良の形態】

【0027】

デジタル信号のデジタル伝送を容易にして、対応する利益を享受するために、一般的にある形態の信号圧縮を採用することが必要である。結果として生じる画像の高い圧縮を実現する一方、画像の高品質が維持されることもまた、重要である。さらに、コンパクトなハードウェア実施のための計算効率が高まる。この計算効率は多くのアプリケーションにおいて重要である。

【0028】

本発明の一実施形態を詳細に説明する前に、本発明は、そのアプリケーションにおいて、構成の詳細におよび以下の記載で述べるまたは、図で図解される部品の配列に限定されないことが理解されるべきである。本発明は他の実施形態を実施することができ、種々の方法で実行される。また、ここで使用される言葉造りおよび専門用語は、記載の目的のためであり、制限と見なされるべきではない。

【0029】

一実施形態の一観点において採用される画像圧縮は、この発明の譲受人に譲渡され、参照することによりここに組み込まれる、1999年11月8日に出願されたシリアル番号第09/436,085の同時係属米国特許出願（発明の名称：「コントラスト感応分散に基づく適応ブロックサイズ画像圧縮」）に開示されるような離散コサイン変換（DCT）に基づく。DCTを利用した画像圧縮および解凍システムは、この発明の譲受人に譲渡され、参照することによりここに組み込まれる、2000年、1月28日に出願されたシリアル番号第09/494,192号の同時係属米国特許出願（発明の名称：「品質に基づく画像圧縮」）に記載されている。一般に、デジタル領域において処理される画像は、 $N \times N$ のサイズの重畳しないブロックの配列に分割されたピクセルデータから成る。二次元のDCTを、各ブロック上で実行してもよい。二次元のDCTは、以下の関係によって定義される：

【数1】

$$X(k, l) = \frac{\alpha(k)\beta(l)}{\sqrt{N \cdot M}} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(m, n) \cos\left[\frac{(2m+1)\pi k}{2N}\right] \cos\left[\frac{(2n+1)\pi l}{2N}\right]$$

$$, 0 \leq k, l \leq N-1$$

【0030】

但し、

【数2】

$$\alpha(k), \beta(k) = \begin{cases} 1, & \text{if } k = 0 \\ \sqrt{2}, & \text{if } k \neq 0 \end{cases}$$

【0031】

および $x(m, n)$ は $N \times M$ ブロック内のロケーション (m, n) におけるピクセルであり $X(k, l)$ は対応するDCT係数である。

【0032】

ピクセル値が0以上であるので、DCT成分 $X(0,0)$ は常に正で、通常最も多くのエネルギーを有する。実際、代表的な画像の場合、大部分の変換エネルギーは、成分 $(0,0)$ 周辺に集中する。このエネルギー圧縮特性は、DCT特性をかくも魅力的な圧縮方法にしているところのものである。

【0033】

画像圧縮技術は、コントラスト適応符号化を利用して、さらなるビットレートの低減を得る。大部分の自然の画像は、相対的にゆっくりと変化する平らな領域、および物体境界および高コントラストテクスチャのような人通りの多い場所から構成される。コントラスト適応符号化スキームは、より多くのビットを人通りの多い場所に割当てより少ないビットをより少ない人通りの多い場所に割当てることにより、この要因を利用する。

【0034】

コントラスト適応方法は、フレーム間符号化（時空間的処理）の代わりに、フレーム内符号化（空間処理）を利用する。フレーム間符号化は、本質的に、より多くの複雑の処理回路に加えて複数のフレームバッファを必要とする。多くのアプリケーションにおいて、複雑さを低減することは、実際の実施のために必要である。フレーム間符号化は、また時空間的符号化スキームの機能が衰えて、その動作が悪化した状況で有効である。例えば、毎秒24コマの映画は、機械的シャッターによる積分時間が相対的に短いので、このカテゴリに入れることができる。短い積分時間は、より高い度合いの時間的エリアシングを可能にする。高速に移動すると完結的になるので、フレームとフレームの相関関係の仮定がこわれる。フレーム内符号化はまた、50Hzおよび60Hzの電力線周波数が含まれる時に、標準化するのがより容易である。テレビ放送は、現在50Hzか60Hzで信号を送る。デジタル手法である、フレーム内スキームの使用は50Hzおよび60Hzの両方に適合する。あるいは、空間分解能に対してフレームレートをトレードオフすることにより毎秒24コマの映画にさえも適用できる。

【0035】

画像処理の目的のために、重量しないブロックの配列に分割されるピクセルデータにDCT動作が実行される。ブロックサイズは $N \times N$ のサイズであるとして本願明細書において論じられるけれども、種々のブロックサイズを用いても良いことが想像される点に留意すべきである。例えば、 $N \times M$ のブロックサイズは、 M が N より大きいまたは未満であった、 N および M の両方が整数である場合に、利用してもよい。他の重要な観点は、ブロックが、 $N/i \times N/j$ 、 $N/i \times N/j$ 、 $N/i \times M/j$ （但し、 i および j は整数）等のような少なくとも1つのレベルの副ブロックに分割可能である。さらに、本願明細書において議論される例示ブロックサイズは、DCT係数の対応するブロックおよび副ブロックを有する 16×16 のピクセルブロックである。偶数または奇数の両方の整数値、例えば 9×9 のような種々他の整数を使用しても良いことがさらに想像される。

【0036】

図1および図2は構成可能なシリアルライザーの概念を組み込んだ画像処理システム100を図解する。画像処理システム100は受信したビデオ信号を圧縮するエンコーダ104を含む。圧縮信号は、伝送路または物理的な媒体108を使用して送られ、デコーダ112によって受信される。デコーダ112は、受信した符号化されたデータを画像サンプルに復号し、次に、それを表示してもよい。

【0037】

一般に、画像は処理のためにピクセルのブロックに分割される。イメージは、加工度のための画素のブロックに分割される。カラー信号は、RGB-YC1C2変換器116を使用して、RGB空間からYC1C2空間に変換してもよい。この場合、Yは輝度成分、または光度成分であり、C1およびC2はクロミナンス成分または色成分である。色に対する眼の空間感度が低いので、多くのシステムは水平方向および垂直方向にC1成分およびC2成分を4だけサブサンプルする。しかしながら、サブサンプリングは必要でない。4:4:4フォーマットとして知られる、完全分解能はデジタル映画をカバーするもの

と言われているようないくつかのアプリケーションにおいて非常に有効または必要かもしれない。2つの可能なY C 1 C 2表示はY I Q表示およびY U V表示であり、両方は技術的によく知られている。また、Y C b C rとして知られているY U V表示の変化を採用することも可能である。これを、さらに、奇数成分と偶数成分に分解してもよい。したがって、実施の形態において、表示Y-偶数、Y-奇数、C b-偶数、C b-奇数、C r-偶数、C r-奇数が使用される。

【0038】

好適実施形態において、偶数および奇数のY、C B、およびC r成分の各々はサブサンプリング無しに処理される。このように、16×16ブロックのピクセルの6つの成分の各々の入力はエンコーダ104に供給される。図示するために、Y偶数成分のためのエンコーダ104が図解されている。同様のエンコーダが、Y-奇数成分および偶数および奇数のC bおよびC r成分のために使用される。エンコーダ104は、ブロックサイズ割当て構成要素120を含み、ブロックサイズ割当て構成要素120はビデオ圧縮の準備のために、ブロックサイズ割当てを実行する。ブロックサイズ割当て構成要素120は、ブロック内の画像の知覚特性に基づいて16×16のブロックのブロック分解を決定する。ブロックサイズ割当ては、各16×16のブロックを、16×16ブロック内の行動に応じて、四部木態様で8×8、4×4、および2×2のようなより小さなブロックにさらに分割する。ブロックサイズ割当て構成要素120は、長さが1ビットと21ビットとの間にあるPQRデータと呼ばれる、四部木データを発生する。従って、ブロックサイズ割当てが、16×16ブロックサイズを分割すると決定するなら、PQRデータのRビットが設定され、4つの分割された8×8ブロックに対応するQデータの4つのさらなるビットが続く。ブロックサイズ割当てが、8×8ブロックのいずれかを分割することを決定するなら、さらに分割された8×8ブロックの各々に対してPデータのさらなるビットが追加される。

【0039】

図3を参照すると、ブロックサイズ割当て構成要素120の動作の詳細を示すフロー図が提供される。ブロックの分散が、ブロックをさらに分割するという決定の測定基準として使われる。ステップ202から始まって、16×16ブロックのピクセルが読まれる。ステップ204において、16×16のブロックの分散が計算される。

【0040】

分散は、次のように計算される：

【数3】

$$\text{var} = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_{i,j}^2 - \left(\frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_{i,j} \right)^2$$

【0041】

但し、N=16であり、x_{i,j}はN×Nブロック内のi行j列のピクセルである。ステップ206において、閾値T16の分散が変更され新しい閾値T'16が与えられ、ブロックの平均値が2つの所定値間にあるなら、そのブロック分散は新しい閾値T'16と比較される。

【0042】

分散16が閾値T16より大きくなければ、ステップ208において、16×16ブロックの開始アドレスが一時的記憶装置に書き込まれ、PQRデータのRビットは0に設定され、16×16ブロックの再分割されないことを示す。次に、アルゴリズムは次の16×16ブロックのピクセルを読む。分散v16が閾値T16より大きければ、ステップ210において、PQRデータのRビットが1に設定され、16×16ブロックが4つの8か

× 8 ブロックに再分割されることを示す。

【 0 0 4 3 】

4 つの 8 × 8 ブロック、すなわち $i = 1 : 4$ は、ステップ 2 1 2 に示すように、さらなる再分割のために順次考慮される。各 8 × 8 ブロックに対して、分散、 $v 8 i$ がステップ 2 1 4 において、計算される。ステップ 2 1 6 において、最初に、分散閾値 $T 4$ が変更され、新しい閾値 $T' 8$ が与えられ、ブロックの値が 2 つの所定値間にあれば、ブロック分散はこの新しい閾値と比較される。

【 0 0 4 4 】

ステップ 2 1 8 において、分散 $v 8 i$ が閾値 $T 8$ より大きくなければ、8 × 8 ブロックの開始アドレスは一時記憶装置に書き込まれ、対応する Q ビット、 $Q i$ は 0 に設定される。次に、次の 8 × 8 ブロックが処理される。ステップ 2 2 0 において、分散 $v 8 i$ が閾値 $T 8$ より大きければ、対応する Q ビット、 $Q i$ は 1 に設定され、8 × 8 ブロックが 4 つの 4 × 4 ブロックに再分割されることを示す。

【 0 0 4 5 】

4 つの 4 × 4 ブロック $j i = 1 : 4$ は、ステップ 2 2 2 に示すように、さらなる再分割のために、順次考慮される。各 4 × 4 ブロックに対して、分散、 $v 4 i j$ はステップ 2 2 4 において、計算される。ステップ 2 2 6 において、最初に、分散閾値 $T 4$ が変更されて、新しい閾値 $T' 4$ が与えられ、ブロックの平均値が 2 つの所定間であれば、ブロック分散はこの新しい閾値と比較される。

【 0 0 4 6 】

分散 $v 4 i j$ が閾値 $T 4$ より大きくなければ、ステップ 2 2 8 において、4 × 4 ブロックのアドレスが書き込まれ、対応する P ビット、 P は 0 に設定される。次に、次の 4 × 4 ブロックが処理される。分散 $v 4 i j$ が閾値 $T 4$ より大きければ、ステップ 2 3 0 において、対応する P ビット、 $P i j$ は 1 に設定され、4 × 4 ブロックは 4 つの 2 × 2 ブロックに再分割されることを示す。さらに、4 つの 2 × 2 ブロックのアドレスは一時記憶装置に書き込まれる。

【 0 0 4 7 】

閾値 $T 1 6$ 、 $T 8$ および $T 4$ は所定の定数であってよい。これは、つらい決断として知られる。あるいは、適応的意思決定または柔軟な意思決定を実施してもよい。例えば、柔軟な意志決定は 2 × 2 ブロックの平均ピクセル値に応じて分散のための閾値を変化させる（但し N は 8、4、または 2 の値を取り得る）。従って、平均ピクセル値の関数を閾値と用いてもよい。

【 0 0 4 8 】

説明のために以下の例を考察する。Y 成分のための所定の分散閾値を、それぞれ 1 6 × 1 6 ブロック、8 × 8 ブロック、および 4 × 4 ブロックの場合に、5 0、1 1 0 0、および 8 8 0 とする。言い換えれば、 $T 1 6 = 5 0$ 、 $T 8 = 1 1 0 0$ 、および $T 4 = 8 8 0$ とする。平均値の範囲を 8 0 および 1 0 0 とする。1 6 × 1 6 ブロックの計算された分散を 6 0 と仮定する。6 0 は $T 1 6$ より大きいので、平均値 9 0 は 8 0 と 1 0 0 との間にあり、1 6 × 1 6 ブロックは 4 つの 8 × 8 の副ブロックに再分割される。8 × 8 ブロックの計算された分散が、1 1 8 0、9 3 5、9 8 0、および 1 2 1 0 であると仮定する。8 × 8 ブロックのうちの 2 つは $T 8$ を越える分散を有するので、これらの 2 つのブロックはさらに再分割され、合計 8 つの 4 × 4 の副ブロックを生成する。最後に、8 つの 4 × 4 ブロックの分散が 6 2 0、6 3 0、6 7 0、6 1 0、5 9 0、5 2 5、および 6 9 0 であり、対応する平均値が 9 0、1 2 0、1 1 0、1 1 5 であると仮定する。最初の 4 × 4 ブロックの平均値は範囲 (8 0、1 0 0) に入るので、閾値は、8 8 0 未満である $T' 4 = 2 0 0$ に下がるであろう。従って、この 4 × 4 ブロックは 7 番目の 4 × 4 ブロックと同様に再分割される。

【 0 0 4 9 】

同様の手続を用いて、輝度成分 Y-奇数および色成分 C b および C r に対してブロックサイズを割当てる点に留意する必要がある。色成分は、水平方向、垂直方向、またはその

両方において、大幅に低減してもよい。

【0050】

さらに、ブロックサイズ割当てがトップダウン手法として記述されたけれども、最大のブロック（この例では 16×16 ）が最初に評価され、その代わりにボトムアップ手法を用いても良いことに留意する必要がある。ボトムアップ手法は、最初に最も小さいブロック（この例では、 2×2 ）を評価するであろう。

【0051】

図1に戻ると、PQRデータは選択されたブロックのアドレスとともに、DCT構成要素124に供給される。DCT構成要素124はPQRデータを用いて、選択されたブロックに関して適当なサイズの離散コサイン変換を実行する。選択されたブロックだけがDCT処理を行う必要がある。

【0052】

画像処理システム100はまた、DCTのDCT係数の中で冗長度を低減するためにDQT構成要素128から成る。DCT係数は、各DCTブロックの左上隅にある。DCT係数は、一般に、AC係数に比べて大きい。サイズにおける不一致は、有効な可変長コーダを設計することを困難にさせる。従って、DCT係数の中で冗長度を低減することは利点がある。

【0053】

DQT構成要素128はDCT係数に2-D DCTsを実行し、一度に 2×2 の単位で行なう。 4×4 ブロック内で、 2×2 ブロックから始めて、4つのDCT係数に関して2-D DCTが実行される。この 2×2 DCTは微分四部木変換、または4つのDCT係数のDQTと呼ばれる。次に、 8×8 ブロック内の3つの隣接するDCT係数とともにDQTのDCT係数を用いて次のレベルDQTを計算する。最後に、 16×16 ブロック内の4つの 8×8 ブロックのDCT係数を用いてDQTを計算する。従って、 16×16 ブロックにおいて、1つのDCT係数があり、残りはDCT及びDQTに対応するAC係数である。

【0054】

変換係数（DCTおよびDQTの両方）は量子化のために量子化器に供給される。好適実施形態において、DCT係数は、周波数重み付けするマスク（FWMs）および量子化スケール因子を用いて量子化される。FWMは、入力DCT係数のブロックと同じ次元の周波数重みのテーブルである。周波数重みは、異なるDCT係数に異なる重みを印加する。重みは、人間の視覚または光学システムがより敏感である周波数内容を有する入力サンプルを強調するように設計され、そして、視覚または光学システムがより敏感でない周波数内容をされている。重みは、また見える距離等のような因子に基づいて設計してもよい。

【0055】

重みは、経験によるデータに基づいて選択される。 8×8 DCT係数のための重み付けマスクを設計するための方法は、ISO/IEC JTC1 CD 10918「連続階調静止画像のデジタル圧縮および符号化—パート1：要件およびガイドライン」、国際標準化機構1994、これは参照することにより本願明細書に組み込まれる。一般に、2つのFWMは、1つが輝度成分のためのそして1つがクロミナンス成分のために設計される。ブロックサイズ 2×2 、 4×4 のためのFWMテーブルは大幅に減らすことにより得られ、 16×16 のためのFWMテーブルは、 8×8 ブロックのためのFWMテーブルの補間により得られる。スケールファクタは、量子化された係数の質およびビットレートを制御する。

【0056】

従って、各DCT係数は以下の関係式に従って量子化される。

【数4】

$$\text{DCT}_q(i, j) = \left\lfloor \frac{8 * \text{DCT}(i, j)}{\text{fwm}(i, j) * q} + \frac{1}{2} \right\rfloor$$

【0057】

但し、 $\text{DCT}(i, j)$ は入力 DCT 係数であり、 $\text{fwm}(i, j)$ は周波数重み付けするマスクであり、 q はスケールファクタであり、および $\text{DCT}_q(i, j)$ は量子化された係数である。DCT 係数の符号に応じて、大括弧内の最初の項は端数のない形に切り上げられるかまたは端数が切り捨てられることに留意する必要がある。DQT 係数は、また、適切な重み付けマスクを使用して量子化される。しかしながら、複数のテーブルまたはマスクを使用することができ、Y 成分、Cb 成分、および Cr 成分の各々に印加することができる。

【0058】

次に、AC 値が DC 値から分離され 130、別個に処理される。DC 構成要素に対して、各スライスの最初の DC 成分値が符号化される。従って、各スライスの各次の DC 成分値は、自分自身と自分より先行する DC 成分値との差分として表され、符号化される 134。損失の無い符号化の場合、各スライスの初期 DC 成分値と前記差分は図 6 および図 8 に対して記載したように、ゴロンブ・ライスを用いて符号化される 138。連続する DC 成分値間の差分のためにゴロンブ・ライス符号化を使用することは、DC 成分値の差が 2 辺のある指数分布を有する傾向があるという点で利点がある。次にデータは、一時的にバッファ 142 を用いて記憶され、次に、送信チャネル 108 を介してデコーダ 112 に転送または送信される。

【0059】

図 8 は DC 成分値を符号化するプロセスを図解する。このプロセスは等価的に静止画、（これらに限定されるわけではないが、映画または高品質テレビジョンのような）ビデオ画像、およびオーディオに適用できる。データ 804 の所定のスライスに対して、スライスの第 1 の DC 成分値が検索される 808。次に、第 1 の DC 成分値が符号化される 812。AC 成分値と異なると、DC 成分値は、量子化される必要はない。一実施形態においてもかわらず、使用される。8×8 または 4×4、またはブロックサイズ割当てにより定義されるいずれかの可変ブロックサイズのような固定のサイズのブロックも使用できると予想される。次に、所定のスライスの第 2 または次の、DC 成分値が検索される 816。次に第 2 の DC 成分値が第 1 の DC 成分値と比較され、その差分すなわち残差が符号化される 820。従って、第 2 の DC 成分値は、自分自身と第 1 の値との間の差分として表すしか必要としない。このプロセスは、スライスの各 DC 成分値に対して繰り返される。このように、問合せ 824 はスライス（最後のブロックそれゆえ、最後の DC 値）の終わりに到達したかどうかに関与しなされる。そうでなければ 828、スライスの次の DC 値が検索され 816、プロセスは繰返される。そうであれば、832、次のスライスが検索され 804、このプロセスは、フレームの全てのスライス、およびファイルの全てのフレームが処理されるまで繰返される。

【0060】

DC 成分値の損失の無い符号化の目的は、低分散を有する傾向がある残差値を生成することである。DCT を使用する際に、DC 係数成分値は、最大の画素エネルギーを与える。それゆえ、DC 成分値を量子化しないことによって、残差の分散は、低減される。

【0061】

AC 要素の場合、データのブロックおよび周波数重み付けマスクは次に、量子化器 146 またはスケールファクタ・エレメントにより倍率がかけられる。DCT 係数の量子化は、大多數の DCT 係数をゼロまで低減し、その結果圧縮が生じる。好適実施形態において、平均ビットレートに対応する 32 のスケールファクタがある。MPEG 2 のような他の圧縮方法と異なり、平均ビットレートは目標のビットレートおよびバッファステータスの代

わりに処理される画像の品質に基づいて制御される。

【0062】

さらに圧縮を増大させるために、量子化されたAC係数はシリアルライザ150に供給される。シリアルライザ150は量子化された係数のブロックを走査し、順番に並べられたストリームの量子化された係数を生成する。ジグザグ走査、カラム走査または行走査を採用してもよい。ジグザグ以外のパターン同様、多数の異なるジグザグ走査パターンを選択してもよい。好適な技術は、ジグザグのスキャンのための8×8ブロックサイズを採用する。量子化された係数のジグザグのスキャンは、ゼロ値の大きいランに遭遇する機会を改良する。このゼロのランは、本質的に減少する確率を有し、ハフマン符号を使用して、効率的に符号化してもよい。

【0063】

順番に並べられ、量子化されたAC係数のストリームは、可変長コーダーに供給される。AC成分値は、ハフマン符号またはゴロンブーライズ符号化のいずれかを用いて符号化してもよい。DC成分値の場合、ゴロンブーライズ符号化が利用される。ランレングスコーダーは、係数をゼロの係数と非ゼロの係数に分けるが、詳細は図6に対して記載される。一実施形態において、ゴロンブーライズ符号化が利用される。ゴロンブーライズ符号化は指数分布を有する負でない整数を符号化する際に効率的である。ゴロンブ符号を用いることにより、指数的に分布される変数に対してより短いレングスコードを供給する際の圧縮のためにより最適である。

【0064】

ゴロンブ符号化ランレングスにおいて、ゴロンブ符号は負でない整数mによりパラメータ化される。例えば、パラメータmを与えられ、正の整数nのゴロンブ符号化は、単項式符号の n/m の商により表されその後に変更されたバイナリコードにより表される剰余が続く。この剰余は、剰余が

【数5】

$$2^{\lceil \log_2 m \rceil} - m$$

【0065】

以下ならば、

【数6】

$$\lfloor \log_2 m \rfloor$$

【0066】

ビット長であり、そうでなければ

【数7】

$$\lfloor \log_2 m \rfloor$$

【0067】

ビット長である。ゴロンブーライズ符号化は、パラメータmが $m = 2^k$ として表される場合のゴロンブ符号化の特別な場合である。そのような場合、 n/m の商は整数nのバイナリ表示をkビットだけ右にシフトすることにより得られ、 n/m の残りはnの最小のkビットにより表される。従って、ゴロンブーライズ符号は2つの連結である。ゴロンブーライズ符号化は、

【数8】

$$P_{\alpha}(x) = c\alpha^{|x|} \quad (1)$$

【0068】

により与えられる二辺の幾何学的な（指数関数）分布を有した正および負の整数を符号化するために使用することができる。

【0069】

(1)において、 α は x の確率の減少を特徴づけるパラメータであり、 c は正規化定数である。 $P_{\alpha}(x)$ は単調であるので、整数値の系列は、

【数9】

$$P_{\alpha}(x_1 = 0) \geq P_{\alpha}(x_1 = -1) \geq P_{\alpha}(x_1 = +1) \geq P_{\alpha}(x_1 = -2) \geq \dots \quad (2)$$

【0070】

を満足することがわかる。

【0071】

図4A、4B、4Cおよび5A、5B、5Cに図解するように、量子化された係数マトリクスにおけるゼロランおよび振幅は、指数分布を有する。これらの図に図解される分布は、実像からのデータに基づく。図4Aは、相対周波数に対するゼロランレングスのY成分分布400を図解する。同様に、図4Bおよび4Cは、それぞれ相対周波数410および420に対するゼロランレングスのCbおよびCr成分分布を図解する。図5Aは、相対周波数に対する振幅サイズのY成分分布500を図解する。同様に、図5bおよび5cはそれぞれ、相対周波数に対する振幅サイズのCb成分分布およびCr成分分布を図解する。図5A、5B、および5Cにおいて、プロットは、DCT係数のサイズの分布を表すことに留意する必要がある。各サイズは、係数値のレンジを表す。例えば、4のサイズ値は、レンジ{-15, -14, ..., -8, 8, ..., 14, 15}、合計16の値を有する。同様に、10のサイズ値は、レンジ{-1023, -1022, ..., -512, 512, ..., 1022, 1023}、合計1024の値を有する。ランレングスと振幅サイズの両方が指数分布を有することは図4A、4B、4C、5A、5Bおよび5Cからわかる。振幅の実際の分布は、以下の式(3)に適合するように示すことができる：

【数10】

$$P(X_{k,1}) = \frac{\sqrt{2\lambda}}{2} \exp\{-\sqrt{2\lambda}|X_{k,1}|\}, k, 1 \neq 0 \quad (3)$$

【0072】

(3)において、 $X_{k,1}$ はそれぞれ、垂直次元および水平次元における周波数 k および1に対応するDCT係数を表し、平均値 μ_x は

【数11】

$$\mu_x = \frac{1}{\sqrt{2\lambda}}$$

【0073】

を表し、分散

【数12】

$$\sigma^2_x$$

【0074】

は、

【数13】

$$\sigma^2_x = \frac{1}{2\lambda}$$

【0075】

を表す。したがって、記載した方法で、ゴロンブライシ符号化を使用することはDCTsにおいてデータを処理する際に、より最適である。

【0076】

以下は、画像データの圧縮に対して記載されるけれども、実施形態は、等価的にオーディオデータを圧縮する実施形態に適用可能である。画像データを圧縮する際に、例えば、画像信号またはビデオ信号は、線形のまたはログ符号化されたピクセル値を有したRGB成分、またはYIQ成分またはYCbCr成分であってもよい。

【0077】

図6はゼロ係数および非ゼロ係数のプロセス600を図解する。DCTマトリクスがキャンセルされると、ゼロ係数および非ゼロ係数が別個に処理され、分離される604。ゼロデータの場合、ゼロランのレンジスが決定される608。ランレンジスは正の整数であることに留意する必要がある。例えば、ランレンジスがnであるとわかると、ゴロンブパラメータmが決定される612。実施形態において、ゴロンブパラメータは、ランレンジスの関数として決定される。他の実施の形態において、ゴロンブパラメータ(m)は以下の式(4)により決定される。

【数14】

$$m = \lceil \log_2 n \rceil \quad (4)$$

【0078】

状況に応じて、ランレンジスの長さおよび関連するゴロンブパラメータはカウンタまたはレジスタによりカウントされる616。nのラベルがつけられたゼロのランレンジスを符号化するために、商が符号化される620。一実施形態において、商はゼロのランレンジスの関数およびゴロンブパラメータとして決定される。他の実施の形態において、商(Q)は、以下の式(5)により決定される。

【数15】

$$Q = \lfloor n / 2^m \rfloor \quad (5)$$

【0079】

一実施形態において、商Qは実施例において、商Qは単項コードで符号化される。これはQ+1ビットを必要とする。次に、剰余が符号化される624。一実施形態において、剰余はランレンジスおよび商の関数として符号化される。他の実施の形態において、剰余(R)は、以下の式(6)を用いて決定される。

【数16】

$$R = n - 2^m Q$$

(6)

【0080】

他の実施形態において、剰余 R は m ビットバイナリコードで符号化される。その後、商 Q および剰余 R が決定され、 Q および R に対する符号が連結され 628、 n のラベルが付けられたランレングスのための全体のコードを表す。

【0081】

非ゼロ係数もゴロンブライスをういて符号化される。係数振幅は正または負でありえるので、サインビットを用いる必要があり、与えられた振幅の絶対値を符号化する必要がある。非ゼロ係数の振幅が x であると仮定すると、振幅の絶対値およびサインの関数として表しても良い。したがって、振幅は以下の式 (7) を用いて y として表すことができる。

【数17】

$$y = \begin{cases} 2x, & \text{if } x \geq 0 \\ 2|x| - 1, & \text{otherwise} \end{cases} \quad (7)$$

【0082】

従って、非ゼロ係数の値はカウンタまたはレジスタにより最速にカウントされる 632。次に、振幅がゼロより大きいまたは等しいかが決定される。そうであるなら、与えられた値の 2 倍の値として符号化される 640。そうでない場合には、絶対値の 2 倍の値より 1 少ない値として符号化される 644。他のマッピングスキームを使用してよいことが予期される。キーポイントは、値の符号を区別する余分なビットが必要ないということである。

【0083】

式 (7) により表される振幅を符号化することにより、 x の正の値が偶数の整数となり、負の値が奇数の整数となる。さらに、このマッピングは (2) の x の確率割当てを保存する。式 (7) で図示したように、符号化の利点は、正の数および負の数を表すために符号ビットを用いて回避することを可能にする。マッピングが成された後、 y はゼロランに対して成されたと同様の方法で符号化される。全ての係数が現在のブロックにおいて走査されるまで、手続は続けられる。

【0084】

この発明の実施形態は、式 (1) 乃至 (7) の関数として係数およびランレングスの値を決定することであるけれども、式 (1) 乃至 (7) のものを使用する必要はない。画像データおよびオーディオデータのより効率的な圧縮を可能にすることは、ゴロンブライスコードおよび DCT 係数の指数分布の利用である。

【0085】

符号化の後のゼロランは、ゼロラン振幅と区別できないので、固定長の特別な前につけるコードを用いて最初のゼロランの発生をマークすることが必要かもしれない。振幅が非ゼロになった後では、ブロック内がすべてゼロになることはよくあることである。そのような場合、ゴロンブライスコードよりもエンドオブブロック (EOB) コードと呼ばれるコードを使用することはより効率的かもしれない。EOB コードは、また一方、状況に応じて特定の固定長コードである。

【0086】

式 (1) または (3) によれば、DCT 係数マトリクス内の振幅またはランレングスの確率分布は α 又は λ によりパラメータ化される。言外の意味は、特定の DCT 係数ブロックが生じるという状況で符号化効率が改良されるかもしれないということである。従って

、関心の質を符号化するために、適当なゴロンブーライスパラメータを使用してもよい。一実施形態において、それぞれの累積値およびそのような値が生じる対応する回数を計算するために、各ランレングスおよび振幅サイズ値に対してカウンタまたはレジスタが使用される。例えば、累積値を記憶するためのレジスタおよび蓄積されるエレメントの数が、それぞれ、 R_{r1} および N_{r1} であるなら、以下の式(6)をライスーゴロンブパラメータとして用いて、ランレングスを符号化してもよい。

【数18】

$$\left\lceil \log_2 \frac{R_{r1}}{N_{r1}} \right\rceil \quad (6)$$

【0087】

同様の手続を、振幅に対して使用してもよい。

【0088】

残差ピクセルは、最初に、ABSDCTデコーダを用いて圧縮されたデータを解凍し、次に、それをオリジナルデータから減算することにより発生される。残差ダイナミックレンジが小さければ小さいほど、圧縮は高くなる。圧縮はブロックベースなので、残差もブロックベースで発生される。残差ピクセルは両面のある指数分布であって、通常ゼロに中心がある指数分布を有することはよく知られた事実である。ゴロンブーライスコードはそのようなデータに対してより最適であるので、ゴロンブーライスコード化手続きを用いて残差データを圧縮する。しかしながら、符号化すべきランレングスが無いので、特別のコードは必要としない。さらに、EOBコードの必要がない。従って、圧縮されたデータは2つの成分から成る。一方は、損失のある圧縮器からの成分であり、他方は、損失の無い圧縮器からの成分である。

【0089】

動作系列を符号化するとき、一時的な相関性を利用する便宜も得ることができる。一時的な相関性を完全に利用するために、ピクセル置換は最初により推定され、次に、残差ピクセルを得るために動作補償予測が実行される。ABSDCTは順応性のあるブロックサイズ符号化を実行するので、ブロックサイズ情報は、動作による置換の手段として代替的に使用してもよい。さらに簡単にするために、場面変更検出は使用しない。その代わり、系列内の各フレームに対して、最初にイントラフレーム圧縮されたデータが得られる。次に、現在のフレームDCTおよび以前のフレームDCTとの間の差分がブロック単位で発生される。これは、参照することにより本願明細書に組み込まれる、2001年6月7日に出願された米国特許出願シリアル番号第09/877,578にさらに詳細に記載されている。DCT領域のこれらの残差は、ハフマン符号化手続きおよびゴロンブーライスコード化手続きの両方を用いて符号化される。従って、最終の圧縮された出力はフレームあたり最小のビット数を使用する圧縮された出力に対応する。

【0090】

損失の無い圧縮アルゴリズムは、損失の無い部分を取り除くことにより、再度目的を達せ、一方のコードを他方のコードに変換するのに好適であるハイブリッドスキームである。従って、ABSDCTを用いることは、空間領域内のピクセル相関を最大化し、予測スキームにおいて使用された残差ピクセルよりも低い分散を有する残差ピクセルを生じる。全体的なシステムの損失のある部分によって、ユーザは、フレーム間処理の助けを求める必要が無くして、分布目的のために必要な品質およびデータレートを得ることができる。これはデジタル映画アプリケーションのために配布されるプログラムにおいて、特に重要である。何故なら、圧縮された部材の損失のある部分は、配布の際により高いレベルの品質を必要とするからである。

【0091】

図9はハイブリッドの損失の無い符号化装置を図解する。図10はそのような装置上で動作してもよいプロセスを図解する。オリジナルデジタル情報904は記憶装置上に常駐し、または送信される。図9のエLEMENTの多くは、図1および図2に対してより詳細に記載されている。データのフレームは、ブロックサイズ割当てELEMENT912、DCT/DQT変換ELEMENT916および量子化器920で構成される圧縮器908に送られる。DCT/DQTがデータ上で実行された後、そのデータは周波数領域に変換される。1つの出力922において、そのデータは量子化器920により量子化され、出力924に転送される、これは、記憶装置および/またはスイッチングで構成してもよい。上述した処理のすべては、フレーム内ベースである。上記した処理の全ては、フレーム内に基づく。

【0092】

量子化器の出力はまた解凍器928に転送される。解凍器928は圧縮器のプロセスを元に戻し、BSAにより定義されたPQRデータの知識と共に、逆量子化器932およびIDQT/IDCT936を通過する。解凍器940の結果は減算器944に供給され、そこで、オリジナルと比較される。減算器944は、各ブロックに対して未圧縮ピクセルおよび圧縮されたおよび解凍されたピクセルとの間の差分とした残差画像を計算する差分器のような、種々のELEMENTであってよい。さらに、差分器は、条件的なフレーム間符号化のために各ブロックに対してDCT領域内の残差を得るようにしてもよい。解凍されたデータとオリジナルデータとの間の比較の結果948がピクセル残差ファイルである。すなわち、結果948は、圧縮されたおよび解凍されたデータにより経験された損失を示している。従って、オリジナルデータは結果と組み合わせられた出力922に等しい。次に、結果948は順番に並べられ、ハフマンおよび/またはゴロンブーライズ符号器956に送られ、第2出力960として供給される。ハフマンおよび/またはゴロンブーライズエンコーダ956は、ゴロンブーライズ符号化を使用している残差画素を符号化する一種のエントロピー・エンコーダであってもよい。意思決定は、各々のフレームのための最小のビットに基づいて、フレーム間かフレーム間かを使用すべきかどうかなされる。残差のゴロンブーライズ符号化の使用は、システムのより高い全体的な圧縮比に導く。

【0093】

従って、損失の無いフレーム間出力は、損失のある高品質の画像ファイル(922またはA)および残りのファイル(960またはC)の組合せまたはハイブリッドである。

【0094】

フレーム間符号化も利用してもよい。量子化器の出力は、BSAの知識と共に、記憶装置964に転送される。1フレーム分のデータを集めると、減算器966は記憶したフレームを次のフレーム968と比較する。その差分はDCT残差970を生じる。この残差は次に順番に並べられ、および/またはゴロンブーライズ符号化され974、第3出力データセット976を出力924に供給する。従って、BおよびCのフレーム間の損失の無いファイルは、コンパイルされる。従って、組合せ(A+CまたはB+C)のいずれかをサイズの考察に基づいて選択してもよい。さらに、純粋なフレーム内出力は、編集目的のために望ましいかもしれない。

【0095】

図1に戻ると、エンコーダ104により発生された圧縮された画像信号は一時的にバッファ142を用いて記憶してもよく、その次に、送信チャネル108を用いてデコーダ112に送信してもよい。送信チャネル108は磁気または光学の記憶装置のような物理的な媒体であってもよく、または、有線または無線の輸送処理または装置であってもよい。ブロックサイズ割当て情報を含むPQRデータもデコーダ112(図2)に供給される。デコーダ112はバッファ164および可変長デコーダ168を含む。可変長デコーダ168は、ランゲルス値および非ゼロ値をデコードする。可変長デコーダ168は同様に動作するが、図6に記載した方法と反対の方法で動作する。

【0096】

可変長デコーダ168の出力は、採用される走査スキームに従って係数を順序づける逆

シリアルライザ 172 に供給される。例えば、ジグザグ走査、垂直走査、および水平走査の混合が使用されるなら、逆シリアルライザ 172 は、採用された走査のタイプの知識を用いて、その係数を再順序付けするであろう。逆シリアルライザ 172 は、PQR データを受信し、復号係数ブロックへの係数の適切な順序付けを支援する。

【0097】

復号ブロックは、量子化器スケールファクタおよび周波数重み付けマスクの使用により処理を元に戻すために、逆量子化器 174 に供給される。

【0098】

微分四部木変換が適用されたなら、次に、係数ブロックは IDQT エLEMENT 186 に供給され、その後 IDCT エLEMENT 190 が続く。そうでなければ、係数ブロックは、直接 IDCT エLEMENT 190 に供給される。IDQT エLEMENT および IDCT エLEMENT 190 は、係数を逆変換し、ピクセルデータのブロックを生成する。従って、ピクセルデータは、補間され、RGB フォームに変換されなければならないかもしれない、そして、次に、将来の表示のために記憶される。

【0099】

図 7 はゴロンブライズ符号化のための装置を図解する。図 7 の装置は、望ましくは、図 6 に対して記載されたプロセスを実施する。決定器 704 はランレングスおよびゴロンブパラメータ (m) を決定する。状況に応じて、各ランレングスおよび振幅サイズ値に対してカウンタまたはレジスタ 708 が使用され、それぞれの累積値およびそのような値が生じる対応する回数を計算する。エンコーダ 712 は、ランレングスおよびゴロンブパラメータの関数として商 (Q) を符号化する。エンコーダ 712 は、また、ランレングス、ゴロンブパラメータおよび商の関数とし、剰余 (R) を符号化する。他の実施形態において、エンコーダ 712 は、非ゼロデータ値および非ゼロデータ値の符号の関数として非ゼロデータを符号化する。連結器 716 を用いて Q 値を R 値と連結する。

【0100】

例として、種々の実例となる論理ブロック、フローチャート、および本願明細書に開示した実施形態に関連して記載したステップは、特定用途向け集積回路 (ASIC)、プログラマブルロジックデバイス、ディスクリートゲートまたはトランジスタロジック、例えばレジスタおよび FIFO のようなディスクリートハードウェアコンポーネント、一組のファームウェア命令を実行するプロセッサ、いかなる一般的なプログラマブルソフトウェアおよびプロセッサ、またはそれらのいずれかの組合せを用いて実施、または実行してもよい。プロセッサは、有利にはマイクロプロセッサであってよいが、他の形態では、プロセッサ、いかなる一般的なプロセッサ、コントローラ、マイクロコントローラまたはステートマシンであってよい。ソフトウェアは、RAM メモリー、フラッシュメモリー、ROM メモリー、レジスタ、ハードディスク、リムーバブル・ディスク、CD-ROM、DVD-ROM または技術的に知られている他のいずれかの形態の記憶媒体に常駐することができる。

【0101】

好適実施形態の上述の記述は、技術に熟達したいかなる人もがこの発明を製作し、または使用することができることを可能にする。これらの実施例に対する種々の変更は、当業者には容易に、明白であり、本願明細書において定義される一般的な原理は発明の能力を用いずに他の実施形態に適用してもよい。従って、本発明は本願明細書において示される実施形態に限定されることを意図したものではなく、本願明細書において開示される原理および新規な特徴に一致する最も広い範囲が許容されるべきである。

【0102】

本発明の他の特徴および利点は以下のクレームに記載される。

【図面の簡単な説明】

【0103】

【図 1】図 1 は、画像圧縮および処理システムのエンコーダ部分のブロック図である。

【図 2】図 2 は、画像圧縮および処理システムのデコーダ部分のブロック図である。

【図 3】図 3 は、分散ベースのブロックサイズ割当てに關係している処理ステップを例示するフローチャートである。

【図 4 A】図 4 A は、DCT 係数マトリクスの Y 成分ランレングスの指数分布を図解する。

【図 4 B】図 4 B は、DCT 係数マトリクスの C b 成分ランレングスの指数分布を図解する。

【図 4 C】図 4 C は、DCT 係数マトリクスの C r 成分ランレングスの指数分布を図解する。

【5 A】図 5 A は、Y 成分の振幅サイズの指数分布または D C T 係数マトリクスにおける Y 成分の振幅サイズの指数分布を図解する。

【図 5 B】図 5 B は、C b 成分の振幅サイズの指数分布または D C T 係数マトリクスの C b 成分の振幅サイズの指数分布を図解する。

【図 5 C】図 5 C は、C r 成分の振幅サイズの指数分布または DCT 係数マトリクスにおける C r 成分の振幅サイズの指数分布を図解する。

【図 6】図 6 は、ゴロンプーライズ符号化プロセスを図解する。

【図 7】図 7 は、ゴロンプーライズ符号化のための装置を図解する。

【図 8】図 8 は、D C 成分値を符号化するプロセスを図解する。

【図 9】図 9 は、損失の無い圧縮のための装置を図解する。

【図 1 0】図 1 0 は、ハイブリッドの損失の無い圧縮の方法を図解する。